

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN BIOPHYSIQUE ET BIOLOGIE CELLULAIRES

PAR
MARIE-EDITH NEPVEU-TRAVERSY

MISE EN ÉVIDENCE DE BACTÉRIES DANS LE SANG DE SUJETS SAINS À
L'AIDE DE TECHNOLOGIES BIOMOLÉCULAIRES

JANVIER 2009

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

REMERCIEMENTS

Je voudrais, premièrement, remercier ceux qui m'ont donné la chance de travailler sur ce projet de maîtrise :

Le Dr. Marc Sirois

Le Dr. Eric Asselin.

Je voudrais également remercier les gens avec lesquels j'ai travaillé tout au long de ma maîtrise :

L'équipe du laboratoire du Dr. Marc Sirois : Annet Lavallée, Valérie Ouellet,
Karine de Carufel et Jean-Francois Vermette

L'équipe du laboratoire du Dr. Éric Asselin

L'équipe du laboratoire du Dr. Carlos Reyes-Moreno

L'équipe du laboratoire du Dr. Lionel Berthoux

Je voudrais également remercier un ami qui m'a supporté tout au long de ma rédaction : Nicolas Laloux.

Merci à vous tous!

IMPLICATION DES AUTEURS DANS L'ARTICLE SCIENTIFIQUE

J'ai réalisé la totalité des expériences mentionnées dans ce document et j'ai procédé à la rédaction de l'article scientifique et ce, sous la supervision de mon directeur de recherche, le Dr. Marc Sirois et de mon co-directeur, le Dr. Éric Asselin.

Le Dr. Maurice Boissinot, du centre de recherche en infectiologie de l'Université Laval, s'est impliqué particulièrement dans les analyses bioinformatiques.

L'article sera soumis au comité de sélection du journal BMC Microbiology pour publication.

RÉSUMÉ

L'avènement de la biologie moléculaire a permis de mettre en évidence et caractériser plusieurs espèces bactériennes autrefois inconnues et ce, en raison de la difficulté à étudier des bactéries ne pouvant croître en laboratoire. Seulement un faible pourcentage (<1%) de bactéries possède la capacité de croître sur un milieu de culture. Des technologies, telle que la réaction de polymérisation en chaîne, ont facilité la détection et la caractérisation de bactéries dans divers environnements. Comme par exemple l'estomac, que l'on croyait quasi-stérile, en raison de son pH acide, contient plus de 128 phylotypes bactériens. Différentes microflore ont été caractérisées à l'aide de cette technologie et ces microflore se sont avérées beaucoup plus nombreuses et complexes que l'on pensait. Étant donné que chez l'Homme, les seuls sites anatomiques considérés stériles sont les organes internes, l'idée d'y retrouver une microflore a été prise en considération. Deux études ont porté sur ce sujet et elles remettent effectivement en cause la stérilité du sang. De plus, différentes recherches ont démontré que plusieurs transfusions de sang, provenant de donneurs sains, ont causé des infections chez les receveurs. Le sang humain serait-il aussi stérile qu'on le croit? Pour y répondre, nous avons utilisé trois technologies permettant de mettre en évidence la présence d'ADN bactérien dans le sang veineux de sujets sains : la PCR conventionnelle, la PCR quantitative et la trappe à promoteur bactérien. Pour démontrer la viabilité des bactéries, la PCR quantitative a été utilisée sur des échantillons prélevés au jour 0 et après 5 jours d'incubation pour démontrer la croissance bactérienne. La PCR quantitative a permis d'obtenir des résultats démontrant une croissance suite à une incubation, mais l'absence d'une courbe standard ne nous permet pas de prendre ces résultats en considération. Une trappe à promoteur bactérien a permis de mettre en évidence de l'ADN bactérien, similaire à 6 espèces bactériennes, dans le sang de 3 sujets sains. De plus, cette technologie a démontré que des séquences d'ADN humain ont la capacité d'interagir avec la machinerie de transcription bactérienne. Fait plutôt surprenant, étant donné que le mécanisme de transcription diffère passablement d'un organisme eucaryote à un organisme procaryote. Ceci nous montre à quel point le génome est très peu connu. Certaines séquences d'ADN possèdent peut être des rôles qui sont inconnus jusqu'à

présent. Pour la détection d'ADN bactérien dans le sang, on peut conclure que le sang humain contient de l'ADN bactérien. Par contre, la provenance reste inconnue : vestige du système immunitaire suite à l'entrée de bactéries lors du prélèvement ou présence d'une microflore dans le sang?

Mots clés : stérilité, microflore, sang, ADN bactérien, promoteur

TABLES DES MATIÈRES

REMERCIEMENTS	II
IMPLICATION DES AUTEURS DANS L'ARTICLE SCIENTIFIQUE.....	III
RÉSUMÉ.....	IV
TABLES DES MATIÈRES.....	VI
LISTE DES TABLEAUX.....	IX
LISTE DES FIGURES	X
LISTE DES SYMBOLES ET DES ABRÉVIATIONS.....	XII

CHAPITRE 1

INTRODUCTION.....	1
1.1 L'OMNIPRÉSENCE DES BACTÉRIES.....	1
1.1.1 <i>Différentes microflores bactériennes chez l'homme</i>	1
1.2 LA STÉRILITÉ DU SANG.....	5
1.2.1 <i>Cas de transmissions de bactéries lors de transfusions de sang provenant de donneurs sains</i>	6
1.3 MÉTHODES DE DÉTECTION ET IDENTIFICATION DES BACTÉRIES À PARTIR D'ÉCHANTILLONS SANGUINS	7
1.3.1 <i>Méthodes conventionnelles</i>	8
1.3.2 <i>Méthodes actuelles (technologies moléculaires)</i>	9
1.4 L'ARN RIBOSOMAL.....	10
1.4.1 <i>Caractéristiques de l'opéron ARN ribosomal (rrn)</i>	11
1.4.2 <i>Gène rrs codant pour la petite sous-unité 16S</i>	12
1.5 MISE AU POINT D'UNE TRAPPE À PROMOTEURS BACTÉRIENS	14
1.5.1 <i>Promoteurs procaryotes VS eucaryotes</i>	15
1.6 TRAVAUX ANTÉRIEURS PORTANT SUR LA DÉTECTION D'ADN BACTÉRIEN DANS LE SANG	15
1.7 OBJECTIFS ET HYPOTHÈSES.....	16

CHAPITRE 2

MATÉRIEL ET MÉTHODES18

2.1	LES CONTRÔLES	18
2.2	PRÉLÈVEMENTS DE SANG SUR DES SUJETS SAINS	18
2.3	EXTRACTION DE L'ADN DU SANG	19
2.3.1	<i>Efficacité de l'extraction</i>	20
2.3.1.1	<i>Migration des produits PCR et photographie</i>	20
2.4	PCR CONVENTIONNELLE	23
2.4.1	<i>Purification des produits PCR</i>	24
2.4.2	<i>Clonage dans le vecteur pDrive</i>	26
2.4.2.1	Préparation des cellules bactériennes compétentes QIAGEN EZ	27
2.4.2.2	Électroporation des ligations dans les cellules compétentes QIAGENEZ	28
2.4.2.3	« Cracking » pour confirmer l'insertion du produit PCR	29
2.4.2.4	Mini-prep pour purifier l'ADN plasmidique (vecteur pDRIVE)	30
2.4.2.5	Préparation des échantillons pour le séquençage	31
2.5	PCR QUANTITATIVE	31
2.6	TRAPPE À PROMOTEURS BACTÉRIENS	33
2.6.1	<i>Construction du vecteur pSP72-GFP</i>	33
2.6.1.1	Digestion et ligation	34
2.6.1.2	Électroporation des ligations dans la bactérie <i>E. coli</i> DH5α	37

CHAPITRE 3

PCR CONVENTIONNELLE POUR LA DÉTECTION D'ADN BACTÉRIEN.....38

3.1	VÉRIFICATION DE L'EFFICACITÉ DE L'EXTRACTION D'ADN	38
3.2	VÉRIFICATION DE L'EFFICACITÉ DES AMORCES	39
3.3	DÉTECTION D'ADN BACTÉRIEN DANS LE SANG PAR RÉACTION PCR ..42	
3.3.1	<i>PCR avec différents volumes d'ADN génomique et d'amorces</i>	42
3.3.2	<i>PCR avec différentes températures d'appariement des amorces</i>	45
3.3.3	<i>PCR avec une augmentation de la durée de dénaturation</i>	46
3.4	CLONAGE DU PRODUIT PCR DANS LE VECTEUR pDRIVE	47
3.4.1	<i>Électroporation de la ligation dans la bactérie QIAGEN EZ</i>	47
3.5	CONCLUSION	49

CHAPITRE 4

PCR QUANTITATIVE POUR LA DÉTECTION D'ADN BACTÉRIEN ET MISE EN ÉVIDENCE DE LA VIABILITÉ.....51

4.1 PCR QUANTITATIVE SUR DES ÉCHANTILLONS PROVENANT DE TUBES SANS ANTI-COAGULANT	51
4.2 PCR QUANTITATIVE SUR DES ÉCHANTILLONS PROVENANT DE TUBES CONTENANT DE L'HÉPARINE ET DU CITRATE (1 SUJET)	53
4.3 PCR QUANTITATIVE SUR DES ÉCHANTILLONS PROVENANT DE TUBES CONTENANT DU CITRATE (4 SUJETS)	55
4.4 PCR QUANTITATIVE SUR DES ÉCHANTILLONS PROVENANT DE TUBES CONTENANT DE L'HÉPARINE (3 SUJETS).....	56
4.5 CONCLUSION	57

CHAPITRE 5

ARTICLE SCIENTIFIQUE58

5.1 INTERACTION OF HUMAN SEQUENCES WITH THE BACTERIAL RNA POLYMERASE IN <i>ESCHERICHIA COLI</i> REVEALED BY PLASMID-VECTOR PROMOTER TRAPPING	58
5.2 RÉSUMÉ.....	59
5.3 ABSTRACT	60
5.4 INTRODUCTION	61
5.5 METHODOLOGY	62
5.5.1 <i>Plasmid construction</i>	62
5.5.2 <i>Enzymatic digestions</i>	63
5.5.3 <i>Transformation and Selection</i>	63
5.5.4 <i>Sequencing and bioinformatics analysis</i>	64
5.6 RESULTS	65
5.6.1 <i>Transformation efficacy</i>	65
5.6.2 <i>Sequencing results</i>	65
5.6.3 <i>Prediction of promoters in cloned sequences</i>	66
5.7 DISCUSSION	67
5.8 REFERENCES	69
5.9 FIGURES LEGENDS	71

CHAPITRE 6

CONCLUSION GÉNÉRALE83

BIBLIOGRAPHIE.....87

ANNEXE A93

ANNEXE B94

LISTE DES TABLEAUX

TABLEAU 2.1 Caractéristiques des amorces b-actine.....	22
TABLEAU 2.2 Réactifs utilisés lors de la réaction PCR avec les amorces b-actine.....	22
TABLEAU 2.3 Paramètres de l'appareil PTC100 (MJ researchs)	22
TABLEAU 2.4 Caractéristiques des amorces utilisées pour le PCR conventionnel	23
TABLEAU 2.5 Réactifs utilisés lors des réactions PCR avec les amorces 16S RNA E. coli.....	25
TABLEAU 2.6 Différents paramètres utilisés lors des réactions PCR effectuées	26
TABLEAU 2.7 Réactifs utilisés pour le clonage des produits PCR dans le vecteur pDRIVE	27
TABLEAU 2.8 Caractéristiques des amorces pour la PCR quantitative	32
TABLEAU 2.9 Différents réactifs et volumes utilisés pour la PCR quantitative.....	32
TABLEAU 2.10 Caractéristiques des amorces utilisées pour le séquençage des fragments d'ADN insérés dans le vecteur pSP72-GFP.....	37
TABLEAU 3.1 Résultats de l'électroporation de la ligation des produits PCR	47
TABLE 5.1 Number of clones on LB amp100 and GFP positives clones after electroporation of ligations	73
TABLE 5.2 <i>BLAST</i> results for cloned sequences (GFP positives clones) similar to bacterial DNA	75
TABLE 5.3 <i>BLAST</i> results for cloned sequences (GFP positives clones) similar to human DNA	76
TABLE 5.4 Predicted promoters with molquest application for cloned sequences	80

LISTE DES FIGURES

FIGURE 1.1 : Différentes microflore présentent normalement chez un être humain [1].	2
FIGURE 1.2 : Carte du gène ARNr16S de la bactérie <i>Escherichia coli</i> présentant les neufs régions hypervariables (soulignées en vert et numérotées de V1 à V9) [36].	14
FIGURE 2.1 : Schéma illustrant les étapes de l'extraction d'ADN génomique du sang avec la trousse QIAamp DNA Blood Mini Kit.	21
FIGURE 2.2 : Vecteur de clonage pDRIVE (source : http://www1.qiagen.com/).	28
FIGURE 2.3 : Carte du vecteur pSP72 (http://www.promega.com/tbs/tb040/tb040.pdf).	34
FIGURE 2.4 : Sites de reconnaissance de l'enzyme BamHI.	35
FIGURE 2.5 : Carte du vecteur pGreenTIR (http://gillnet.lab.nig.ac.jp/~cvector/map/pGreenTIR.gif).	35
FIGURE 2.6 : Sites de reconnaissance de l'enzyme Sau3AI.	36
FIGURE 2.7 : Carte du vecteur pSP72-GFP (construit en laboratoire).	36
FIGURE 3.1 : Photographie du gel représentant la migration des produits PCR obtenus avec les amorces b-actine sur l'ADN génomique extrait du sang au jour 0 et 5.	38
FIGURE 3.2 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique de différentes bactéries avec une température d'appariement des amorces de 50°C.	40
FIGURE 3.3 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique de différentes bactéries avec une température d'appariement des amorces de 48°C.	41
FIGURE 3.4 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique extrait de <i>E. coli</i> 100.	42
FIGURE 3.5 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique (2µl) avec une température d'appariement des amorces de 48°C.	43
FIGURE 3.6 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique (5µl) avec une température d'appariement des amorces de 48°C.	44

FIGURE 3.7 : Photographie du gel représentant la migration des produits sur l'ADN génomique (5µl) avec une température d'appariement des amorces de 45°C.	45
FIGURE 3.8 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique (5µl) avec une durée de dénaturation de 45s.	46
FIGURE 3.9 : Migration du cracking effectué sur 19 clones recombinants.	48
FIGURE 4.1 : PCR quantitative sur 4 extractions provenant de sang contenu dans des tubes sans coagulant de deux sujets au jour 0 et 5.	52
FIGURE 4.2 : PCR quantitative sur 4 extractions provenant de sang contenu dans des tubes avec deux différents coagulants (citrate et héparine) d'un sujet au jour 0 et 5.	53
FIGURE 4.3: PCR quantitative sur 4 extractions provenant de sang contenu dans des tubes avec deux différents coagulants (citrate et héparine) d'un sujet au jour 0 et 5 (essai #2).	54
FIGURE 4.4: PCR quantitative sur 8 extractions provenant de sang contenu dans des tubes contenant du citrate de quatre sujets au jour 0 et 5.	55
FIGURE 4.5: PCR quantitative sur 3 extractions provenant de sang contenu dans des tubes de l'héparine-sodium de trois sujets au jour 0 et 5.	56
FIGURE 5.1: Plasmidic construction with restriction enzyme <i>Eco</i> RI. Red arrows represent enzymatic digestion sites on plasmids pSP72 and pGreenTIR.	71
FIGURE 5.2: GFP cassette showing enzyme restriction sites and Translation Initiation Region (TIR): enhancer, SD (Shine-Dalgarno) sequence and star codon.	72

LISTE DES SYMBOLES ET DES ABRÉVIATIONS

ADN	Acide Désoxyribonucléique
Amp	Ampicilline
<i>App</i>	<i>Actinobacillus pleuropneumoniae</i>
ARNr	Acide Ribonucléique Ribosomal
<i>BLAST</i>	<i>Basic Local Alignment Search Tool</i>
BRE	TFIIB Recognition Element
°C	Celsius
CMV	CytomégaloVirus
CO ₂	Dioxyde de carbone ou gaz carbonique
DMSO	Diméthyl sulfoside
dNTP	Dinucléotide triphosphate
DPE	Downstream Promoter Element
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylen-Diamine-Tetraacetic Acid
Enh	Enhancer
FISH	Fluorescent <i>in situ</i> hybridisation
GFP	Green Fluorescent Protein
INR	Initiator element
IPTG	Isopropyl β-D-triogalactopyranoside
LB	Luria Bertani
MgCl ₂	Dichlorure de magnésium
Pb	Paire de bases
PCR	Polymerase Chain Reaction
RNAP	RNA polymerase
SD	Shine Dalgarno
SPS	Sulphonate Polyanéthol de Sodium
<i>Taq</i>	<i>Thermophilus aquaticus</i>
T _A	Température d'appariement
TIR	Transcritpion Initiation Region

T _M	Melting temperature
TSA	Tryptic Soy Agar
TSB	Tryptic Soy Broth
TSS	Transcription Start Site
X-gal	5-bromo-4-chloro-3-indolyl- β -D-triogalactopyranoside

CHAPITRE I

INTRODUCTION

1.1 L'OMNIPRÉSENCE DES BACTÉRIES

Les bactéries sont ubiquitaires. Elles sont présentes à peu près partout : (a) dans l'eau, (b) dans l'air, (c) dans la terre et (d) elles sont une partie intégrante du corps humain [1]. On les retrouve même dans des milieux aux conditions extrêmes où la majorité des organismes vivants ne pourraient survivre comme: (a) les sources thermales pouvant atteindre des températures de 110°C [2], (b) les eaux profondes sous haute pression [3] ou (c) des eaux hyper-salées comme celles de la Mer Morte [4]. Elles peuvent également vivre dans des conditions plus modérées comme, entre autres, les bactéries qui colonisent les nombreuses microflore de l'être humain [5]. Ceci est, en partie, dû à leur grande capacité d'adaptation à leur environnement. Pour survivre, les bactéries ont développé plusieurs mécanismes de défense comme par exemple: pour résister à des conditions extrêmes [2, 3] ou pour contourner le système immunitaire dans le cas de bactéries pathogènes [6]. Dans le cas des bactéries colonisant les différentes microflore du corps humain, elles n'utilisent pas un mécanisme de défense, mais elles établissent une relation commensale avec les cellules humaines et elles co-évoluent avec les cellules humaines [7]. Le commensalisme se définit ainsi : symbiose dans laquelle un organisme profite de l'association tandis que l'autre n'est ni atteint, ni avantagé [1].

1.1.1 Différentes microflore bactériennes chez l'homme

Le corps humain possède plusieurs microflore indigènes (figure 1.1). Fait intéressant, le nombre total de bactéries (10^{14}) colonisant le corps humain est supérieur au nombre total de cellules de l'organisme (10^{13}) [7]. La répartition des différentes

microflore se résume aux régions du corps suivantes : (a) la peau, (b) le nez et le rhinopharynx, (c) l'oropharynx, (d) le système respiratoire, (e) la cavité buccale, (f) l'œil, (g) l'oreille externe, (h) l'estomac, (i) l'intestin grêle, (j) le côlon (gros intestin) et (k) le système urogénital [1]. Le type de micro-organismes y résidant dépend de la région.

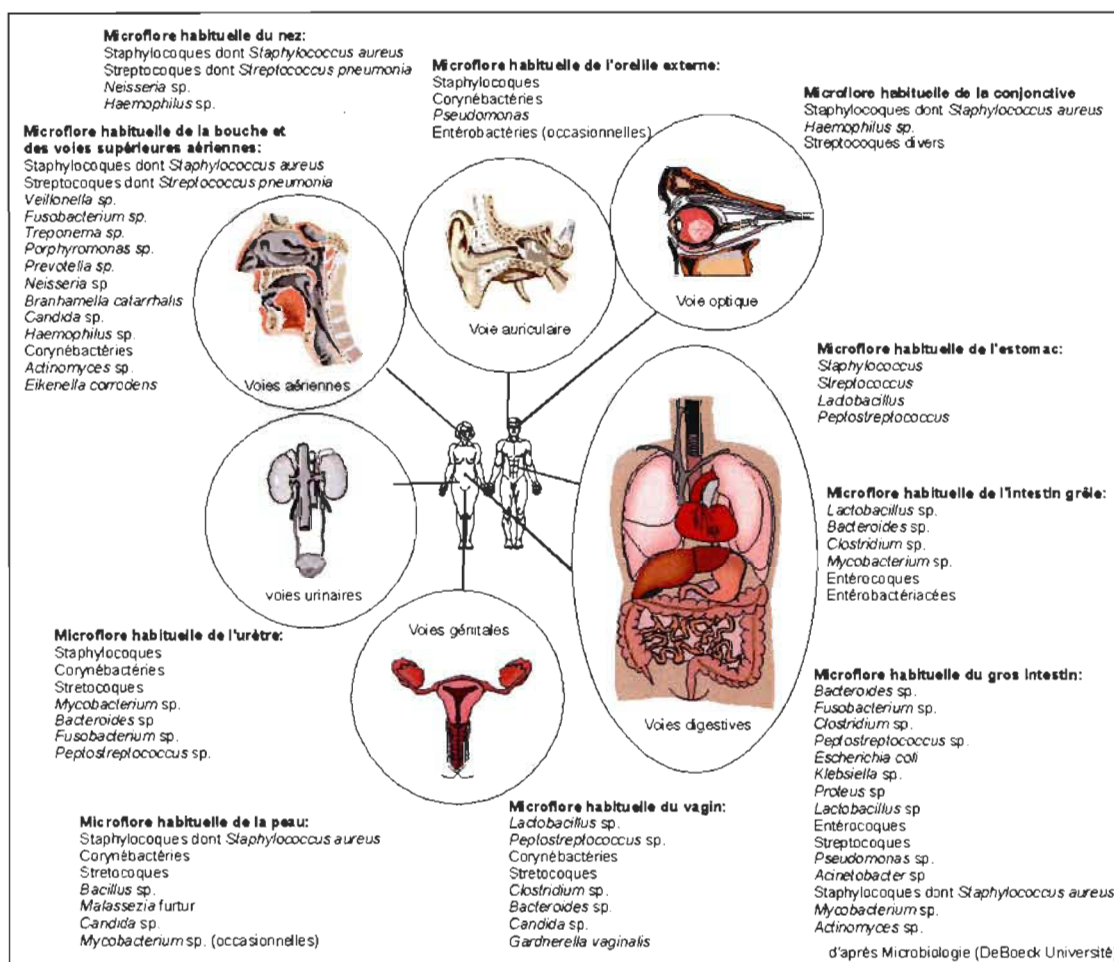


FIGURE 1.1 : Différentes microflore présentes normalement chez un être humain [1].

Pour commencer, la peau possède une anatomie et une physiologie qui varient d'une partie à l'autre du corps. Par exemple, certaines parties du corps ((a) les oreilles, (b) les régions génito-urinaire et anale, (c) le périnée, (d) les paumes) ont une humidité relative assez élevée pour permettre l'existence d'une microflore résidente. Les bactéries Gram-négatives sont généralement présentes dans ces régions plus humides. La majorité

des bactéries sont présentes sur l'épiderme squameux superficiel et sont souvent associées aux glandes sébacées et sudoripares puisqu'elles fournissent (a) de l'eau, (b) des acides aminées, (c) de l'urée, (d) des électrolytes et (e) des acides gras servant d'éléments nutritifs [1]. La division bactérienne qui domine sur toutes les surfaces de la peau est *Proteobacteria*. En plus d'inhiber la majorité des pathogènes, les bactéries indigènes présentes sur la peau agissent dans le métabolisme des protéines, des acides gras et du sébum [8]

La microflore normale du nez se résume principalement aux micro-organismes suivant : *Staphylococcus aureus* et *Staphylococcus epidermidis*. Le rhinopharynx, quant à lui, peut contenir des petites populations (a) de *Streptococcus pneumoniae*, (b) de *Neisseria meningitidis* et (c) d'*Haemophilus influenzae*, mais contrairement aux souches pathogènes, celles-ci sont non capsulées [1].

Pour les bactéries colonisant l'oropharynx, les plus importantes sont les différents streptocoques alpha-hémolytiques ((a) *S. oralis*, (b) *S. milleri*, (c) *S. gordonii*, (d) *S. salivarius*), un nombre élevé de corynébactéries et de petites coques gram-négatifs apparentés à *Neisseria meningitidis*. Les amygdales portent une flore identique à celle de l'oropharynx [1].

Pour ce qui est des systèmes respiratoires supérieur et inférieur (trachée, bronches, bronchioles et alvéoles), ils ne possèdent pas de microflore indigènes puisque la plupart des micro-organismes sont éliminés par le flot incessant de mucus généré par les cellules épithéliales ciliées et l'activité phagocytaire des macrophages alvéolaires [1].

Quant à la cavité buccale, elle contient de nombreux micro-organismes. Par contre pour y résider, ceux-ci doivent être capables de résister à une élimination mécanique en adhérant à différentes surfaces comme les gencives, la langue et les dents [1]. L'existence d'une microflore bactérienne dans la bouche n'est pas un fait nouveau, mais récemment, des études ont permis de démontrer qu'il y avait plusieurs microflore dans la cavité buccale [9, 10]. Par exemple, les bactéries présentes sur la langue ne sont pas

les mêmes que celles sur les côtés latéraux de la langue. En réalité, cette constatation n'est pas surprenante puisque ces différentes structures possèdent différentes fonctions : les côtés latéraux de la langue ont une surface lisse et non kératinisée contrairement au dessus de la langue qui est kératinisée et possède de nombreuses papilles gustatives. Ces différences anatomiques influencent considérablement l'écologie de ces niches et favorisent un environnement microbien différent [9].

L'œil humain contient des bactéries qui tirent un avantage à s'y établir puisqu'il s'agit d'un milieu humide et riche en nutriments [1]. La bactérie qui prédomine est *Staphylococcus epidermidis* suivie de près par *Staphylococcus aureus* [1, 11].

L'oreille externe possède une flore semblable à celle de la peau. De plus, des études mycologiques ont permis de mettre en évidence la présence de mycètes ((a) *Aspergillus*, (b) *Penicillium*, (c) *Candida*, (d) *Saccharomyces*) dans cette microflore [1].

En raison du pH très acide (2.0-3.0) et de la présence de divers agents antibactériens, l'estomac n'a jamais été considéré comme un milieu propice à l'établissement d'une flore bactérienne [1]. Par contre, des études moléculaires récentes, qui ont l'avantage de détecter des bactéries incapables de croître en laboratoire, ont démontré que plus de 128 phylotypes (organisme ayant un certain pourcentage de similarité avec un organisme connu et classifié) étaient présents dans la muqueuse gastrique de 23 sujets. Ceci suggère donc que l'estomac posséderait une microflore beaucoup plus peuplée et diversifiée que l'on croyait [12].

L'intestin grêle comprend trois régions : (a) le duodénum, (b) le jéjunum et (c) l'iléon. Le duodénum contient peu de micro-organismes pour deux raisons : la présence des sucs acides de l'estomac combinés à l'action inhibitrice de la bile et des sécrétions pancréatiques. Dans le jéjunum, on retrouve principalement des coques et des bacilles Gram-positifs. Dans l'iléon, le pH devient plus alcalin permettant ainsi aux bactéries Gram-négatives anaérobies et aux membres des *Enterobacteriaceae* de s'établir pour se rapprocher ainsi de la microflore normale du côlon.

Le côlon est sans aucun doute la région la plus riche en micro-organismes. Les principaux représentants sont des bactéries Gram-négatives anaérobies non sporulantes et des bacilles Gram-positifs sporulants ou non [1]. La flore intestinale diffère considérablement d'un individu à l'autre, dépendamment de la diète de chacun [13].

Le système urogénital supérieur (reins, uretères et vessie) contient peu de micro-organismes. Parfois, quelques bactéries ((a) *Staphylococcus epidermidis*, (b) *Enterococcus faecalis*, (c) *Corynebacterium sp.*) sont retrouvées dans l'urètre chez l'homme et la femme. Par contre, le système génital de la femme adulte possède une microflore complexe qui varie constamment en fonction du cycle menstruel [1]. La santé de la flore vaginale influence grandement le risque de contracter des infections causées par des organismes pathogènes, comme par exemple : des vaginites à levure (*Candida albicans*) ou même la transmission du VIH (virus de l'immunodéficience humaine). Une flore bien équilibrée protège, en partie, contre différentes infections [14].

Les microflores bactériennes sont nombreuses chez l'humain et elles diffèrent d'un individu à l'autre [9]. Elles sont essentielles puisqu'elles jouent un rôle important dans la défense de l'organisme contre les pathogènes. Elles empêchent également l'intrusion de micro-organismes opportunistes susceptibles de causer des maladies. De plus, certaines microflores nourrissent les cellules humaines en agissant en symbiose avec celles-ci, comme c'est le cas des bactéries colonisant le côlon [15]. Par contre, il ne faut pas oublier que ces mêmes microflores peuvent être la cause de certaines maladies, comme par exemple le cas des abcès buccaux causés par *Porphyromonas gingivalis* et *Prevotella oris* [16] ou de certaines parodontites causées aussi par *Porphyromonas gingivalis* [17].

1.2 LA STÉRILITÉ DU SANG

Même si l'Homme possède plusieurs microflores bactériennes peuplées de milliards de bactéries [1], certains organes sont considérés stériles. En réalité, tous les

organes n'ayant pas contact avec le milieu externe ((a) foie, (b) cœur, (c) reins, (d) cerveau, (e) liquide céphalo-rachidien, (f) sang) sont exempts de bactéries ou autres microorganismes. Il existe toutefois des exceptions : infections dues à la présence d'un microorganisme opportuniste ou lorsqu'il y a un bris dans l'intégrité de la membrane, laissant ainsi une porte d'entrée aux bactéries potentiellement pathogènes. Pour le sang, il existe deux exceptions : la bactériémie et la septicémie [1, 18]. La bactériémie se définit par la présence de bactéries dans le sang et elle est souvent associée à un site d'infection. Par exemple, un patient ayant une infection au niveau du tractus urinaire est plus propice à contracter une bactériémie. Pour ce qui est de la septicémie, il s'agit d'une bactériémie accompagnée de signes et symptômes cliniques ((a) fièvre, (b) frissons, (c) hypothermie, (d) hyperventilation et (d) choc septique) suite à la dissémination de la bactérie dans les périphéries [18]. En fait, la stérilité du sang est considérée comme un dogme. À moins d'avoir des signes ou symptômes cliniques dus à une infection, le sang d'individus d'apparence saine est considéré totalement stérile, microbiologiquement [18]. Pourtant, lors de transfusions de sang ou de transplantations d'organes, il y a eu plusieurs infections causées par la présence de bactéries dans le sang de donneurs supposément sains.

1.2.1 Cas de transmissions de bactéries lors de transfusions de sang provenant de donneurs sains

Certaines études, portant sur les agents transmissibles lors de transfusions de sang ou de greffes d'organes, ont démontré qu'il y avait une émergence d'agents transmissibles (de nature virale ou bactérienne) dans le sang de donneurs sains [19, 20]. Actuellement, la mortalité et la morbidité, à la suite de transfusions sanguines, sont principalement dues à des infections causées par des bactéries (septicémies) [21]. Le risque de contracter une infection bactérienne lors de transfusions de sang est très élevé : 1 chance sur 2000. Aux États-Unis, les contaminations bactériennes sont considérées la deuxième cause de mortalité due aux transfusions de sang (après les erreurs administratives) avec des taux de mortalité allant de 1 : 20 000 à 1 : 85 000 [22]. Ceci s'explique principalement par la mise en place de tests permettant de dépister, dans le sang des donneurs, les agents les plus susceptibles d'être transmis par le sang, c'est-à-

dire, les virus ((a) virus de l'immunodéficience humaine (VIH), (b) virus causant les hépatites B et C, (c) cytomégalo virus (CMV)) et les parasites (agent causal de la malaria (*Plasmodium*) et les spirochètes tels que *Leptospira* et *Treponema*) [23]. Ainsi, le taux de mortalité attribuables à ces derniers s'est vu considérablement diminué au cours des dernières années [22, 23]. Donc, dans le but de diminuer le risque de transmission de bactéries potentiellement pathogènes lors de transfusion, plusieurs tests ont été effectués pour déterminer l'origine de ces infections. Dans la plupart des cas, la contamination survient lors de la collecte de sang (désinfection insuffisante de la peau) ou lors de manipulations ultérieures du sang. Par contre, un certain nombre de contaminations serait dû à des bactériémies non détectées chez les donneurs puisqu'ils sont asymptomatiques [21].

En effet, plusieurs études démontrent la présence de bactéries dans le sang de donneurs réguliers, admissibles au don de sang (sans antécédents d'infections virales ou autres) et ayant passé les tests nécessaires au dépistage des nombreux agents transmissibles par le sang. Entre autres, une étude a permis de détecter la présence de *Chlamydia* dans des monocytes (cellules du sang) provenant du sang de donneurs sains [24]. Un cas de contamination des plaquettes sanguines par *Listeria monocytogenes* chez un donneur régulier a également fait l'objet d'une étude [25]. Le développement de nouvelles techniques permettant la détection de bactéries non cultivables en laboratoire est peut-être aussi un facteur contribuant à l'émergence des bactéries lors de transfusions sanguines : auparavant, elles n'étaient pas détectées par les méthodes conventionnelles.

1.3 MÉTHODES DE DÉTECTION ET IDENTIFICATION DES BACTÉRIES À PARTIR D'ÉCHANTILLONS SANGUINS

Pour détecter et identifier les bactéries présentes dans un échantillon de sang, plusieurs techniques sont utilisées. Les méthodes conventionnelles, à la fois laborieuses et pas toujours efficaces, ne permettent de détecter que les bactéries capables de croître sur un milieu artificiel en laboratoire. Étant donné qu'un faible pourcentage ($\leq 1\%$) de bactéries à la capacité de croître sur un milieu de culture, le développement de la

biologie moléculaire a permis de détecter et d'identifier de nombreux microorganismes non cultivables [5]. Puisque les infections du sang sont un problème de santé majeur actuellement, l'avènement de méthodes moléculaires constitue un atout important : la rapidité de détection et l'augmentation de la sensibilité permettent un diagnostic rapide et précis [26].

1.3.1 Méthodes conventionnelles

Il existe plusieurs méthodes conventionnelles pour détecter la présence de bactéries dans le sang : (a) la coloration de Gram, (b) la détection de caractères spécifiques (différence dans la production du gaz carbonique (CO_2) ou du pH, la fermentation du glucose), (c) la détection des antigènes spécifiques et (d) la méthode la plus connue, l'hémoculture. La coloration de Gram est simple, rapide et non dispendieuse, mais possède trois désavantages : (a) pas très sensible, (b) la différence de coloration entre les espèces n'est pas assez grande pour les distinguer entre elles et (c) l'évaluation microscopique dépend de l'expertise. La détection de caractères spécifiques, comme par exemple l'augmentation du CO_2 , est couramment utilisée par contre, même si cette méthode peut prendre quelques heures, certaines espèces mettent jusqu'à 7 jours pour produire du CO_2 . Les tests immunologiques pour la détection des antigènes sont très spécifiques et rapides, mais un nombre limité d'espèces est détectable avec cette technique [27]. Pour ce qui est de l'hémoculture, cette technique consiste à prélever le sang dans un tube contenant un anticoagulant, puisque les microorganismes sont difficiles à isoler à partir de sang coagulé. Le Sulphonate Polyanéthol de Sodium (SPS) est l'anticoagulant recommandé pour les hémocultures. En plus de prévenir la coagulation, celui-ci empêche la phagocytose, inactive le complément et neutralise certains agents antimicrobiens [18].

À la suite du prélèvement, un certain volume de sang estensemencé en bouillon (le plus fréquemment utilisé) ou sur une gélose d'agar. L'incubation s'effectue à des températures se situant entre 35 et 37°C, dépendamment de l'espèce bactérienne. Le temps d'incubation recommandé dépend généralement du type de système utilisé et du

type de microorganisme recherché. La plupart des cultures sont incubées pour une période variant de 5 à 7 jours. Ensuite, les résultats sont interprétés; pour la culture en bouillon, (a) la turbidité, (b) la formation de gaz, (c) la décoloration, (d) l'hémolyse, (e) la coagulation, (f) la présence de colonies visibles et (g) la formation de biofilms peuvent démontrer la croissance bactérienne; pour la culture sur des milieux d'agar, il y a présence de colonies. Par la suite, des sous-cultures sont effectuées pour les hémocultures en bouillon. Quelques gouttes de bouillon sont ensemencées sur milieu d'agar enrichi et incubées à 35-37°C pour une période de 18 à 48 heures [18]. Pour permettre l'identification et la caractérisation des bactéries présentes dans les hémocultures positives, différents tests biochimiques sont effectués [28].

Dans tous les cas, les résultats ne sont pas analysés avant quelques jours. De plus, la plupart des bactéries sont incapables de croître en laboratoire, cette technique s'avère imprécise et limitée pour la détection des bactéries. De plus, la prise d'antibiotiques par le patient, la présence de bactéries fastidieuses (croissance lente) ou intracellulaires [29] empêchent ou ralentissent la détection du microorganisme présent [30].

1.3.2 Méthodes actuelles (technologies moléculaires)

Puisque la méthode conventionnelle de détection des pathogènes dans le sang s'avère limitée et peu rapide, le développement de diverses technologies moléculaires a permis de diagnostiquer rapidement la présence de bactéries pathogènes dans le sang. En réalité, la majorité de ces technologies découle de la réaction de polymérisation en chaîne (PCR), qui est maintenant la plus utilisée et la plus documentée [31, 32]. Le plus grand avantage de cette technologie est qu'elle est indépendante des cultures bactériennes en laboratoire puisqu'elle nécessite seulement l'ADN bactérien pour permettre l'identification [31, 33].

1.3.2.1 La réaction de polymérisation en chaîne

Brièvement, la réaction de polymérisation en chaîne permet d'amplifier une séquence d'acides nucléiques à l'aide d'une réaction enzymatique effectuée par une polymérase d'ADN (acide désoxyribonucléique). Des amorces (oligonucléotides) vont s'apparier aux séquences spécifiques (recherchées) d'acides nucléiques et la polymérase effectuera une copie complémentaire de chaque brin d'ADN. La réaction est répétée plusieurs fois (cycles) pour produire plusieurs copies de la séquence d'acides nucléiques d'intérêt [34]. Cette technique est couramment utilisée pour le diagnostic bactérien dans le sang [29, 30, 35].

Récemment, le développement de la PCR quantitative a permis de faciliter la détection d'ADN bactérien dans les échantillons sanguins. Comparativement à la PCR conventionnelle, la PCR quantitative a plusieurs avantages : (a) elle est rapide, (b) simple d'utilisation, (c) quantitative et (d) le risque de contamination est réduit. La technologie de la PCR quantitative combine l'amplification et la détection des produits amplifiés. La détection s'effectue en temps réel à l'aide de marqueurs fluorescents spécifiques dans le cas de sondes ou non-spécifiques dans le cas de marqueurs intercalants. Ces derniers se lient non spécifiquement aux ADN double-brins produits lors de l'amplification. D'une façon ou d'une autre, la fluorescence est émise par les marqueurs fluorescents à chaque cycle d'amplification et le signal augmente en fonction du nombre de copies d'ADN amplifié [29].

Pour détecter de l'ADN bactérien, la séquence la plus couramment ciblée pour une amplification par PCR est celle correspondant au gène codant pour l'acide ribonucléique ribosomal (ARNr), plus particulièrement celle de la petite sous-unité 16S [36, 37].

1.4 L'ARN RIBOSOMAL

L'ARN ribosomal est essentiel au mécanisme de synthèse des protéines puisqu'il est une constituante du ribosome. Les ribosomes sont le siège physique de la synthèse

des protéines et font partie des plus gros organelles d'une cellule bactérienne. Ils contiennent une cinquantaine de protéines ribosomales et trois types d'ARNr qui forment la structure du ribosome : (a) 16S, (b) 23S et (c) 5S. Ils sont nommés ainsi en fonction du taux de sédimentation lors de la centrifugation. En général, plus l'indice S est élevé, plus grosse est la molécule d'ARN. Leur rôle distinct, lors de la traduction des ARNs messager en protéines, se résume ainsi : l'ARNr 23S est l'enzyme peptidyltransférase qui ajoute les acides aminés à la protéine et l'ARN 16S est directement impliqué dans l'initiation et la terminaison de la traduction.

Les ARNs ribosomaux et les ARNs de transfert sont les ARNs les plus retrouvés (95% de l'ARN total d'une bactérie) dans une cellule bactérienne et ce, pour deux raisons. Dans une bactérie qui est dans la phase exponentielle de croissance, environ la moitié des ARNs synthétisés sont destinés à faire ces deux types d'ARN. De plus, ces ARNs sont plus stables que les ARN messager puisqu'ils ne sont pas dégradés avant une longue période de temps après leur synthèse [38].

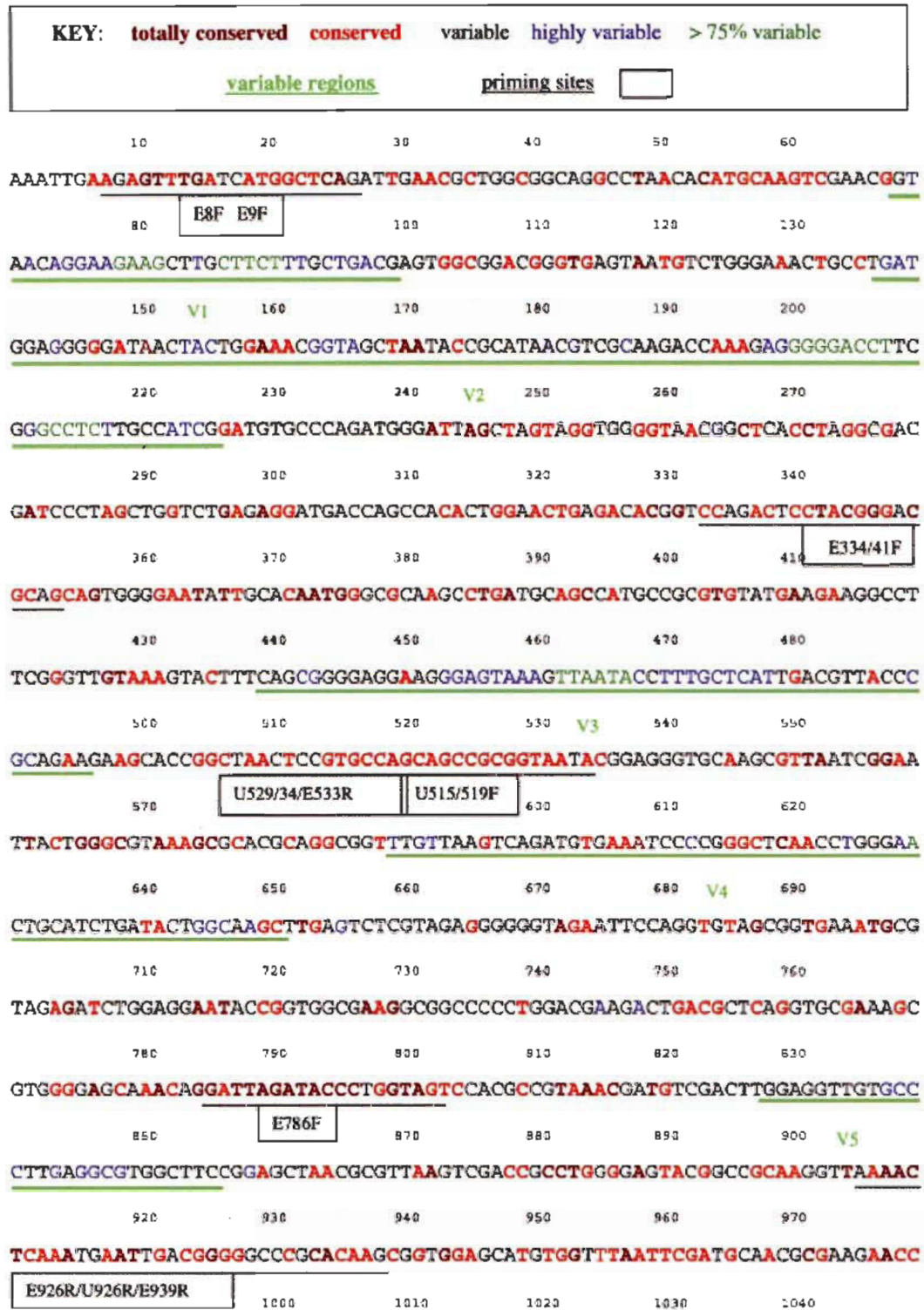
1.4.1 Caractéristiques de l'opéron ARN ribosomal (*rrn*)

En plus d'être associés dans le ribosome, les trois ARNs ribosomaux sont synthétisés à partir d'un seul opéron. Cet opéron contient également des gènes codant pour certains ARNs de transfert, donc, les différentes molécules transcrites à partir de cet opéron sont polycistroniques. Puisque le ribosome est le site de synthèse des protéines, parfois, la cellule peut augmenter sa croissance en augmentant son nombre de ribosomes. En effet, plus vite une bactérie synthétise des protéines plus vite elle croît. Ainsi, plusieurs bactéries contiennent plus d'un opéron (7 à 10) situés dans différents endroits chromosomiques de la cellule [38], comme par exemple la bactérie *Escherichia coli* qui en contient sept [39]. De plus, le nombre total d'opérons par bactérie peut varier selon le nombre de chromosomes lors de la phase exponentielle de croissance [27].

1.4.2 Gène *rrs* codant pour la petite sous-unité 16S

La caractérisation de l'ARNr 16S est devenue possible dans les années 1970 et début des années 1980 et a permis d'élucider la relation évolutive des bactéries pour la première fois. Il est maintenant commun d'identifier la position phylogénétique d'une bactérie inconnue en amplifiant et séquençant son ADNr 16S [40]. Étant donné le risque de contaminations bactériennes lors de transfusions de sang, l'utilisation de l'ARNr 16S a grandement contribué à améliorer la détection de l'ADN bactérien dans les échantillons sanguins [27].

L'utilisation des gènes codant pour l'ARNr 16S, lors de détection de bactéries pour l'identification d'espèces bactériennes ou pour des études taxonomiques, peut s'expliquer par plusieurs avantages. Premièrement, ces gènes sont reconnus pour s'être hautement conservés lors de l'évolution. De plus, comme cité précédemment, les ARNs ribosomiaux sont nombreux dans une cellule bactérienne. Celle-ci peut synthétiser entre 2000 et 50 000 copies d'ARNr 16S (taille de plus ou moins 1540 nucléotides, selon les espèces), dépendamment de son stade de croissance [27]. La présence de neuf régions hypervariables (figure 1.2) permet également une identification précise puisque ces régions sont différentes d'une espèce à l'autre. Ces régions sont entourées de régions hautement conservées. L'utilisation d'amorces ciblant les parties hautement conservées du gène, mais amplifiant une région hypervariable, permet donc une détection et une identification rapide et spécifique d'une bactérie suite à un séquençage [36]. De plus, le gène codant pour l'ARNr 16S est très bien documenté (6800 séquences complètes répertoriées dans les banques publiques) comparativement au gène codant pour l'ARNr 23S (204 séquences complètes répertoriées) [37]. Toutes ces raisons font de l'ADNr 16S une cible idéale pour des études portant sur la détection d'ADN bactérien.



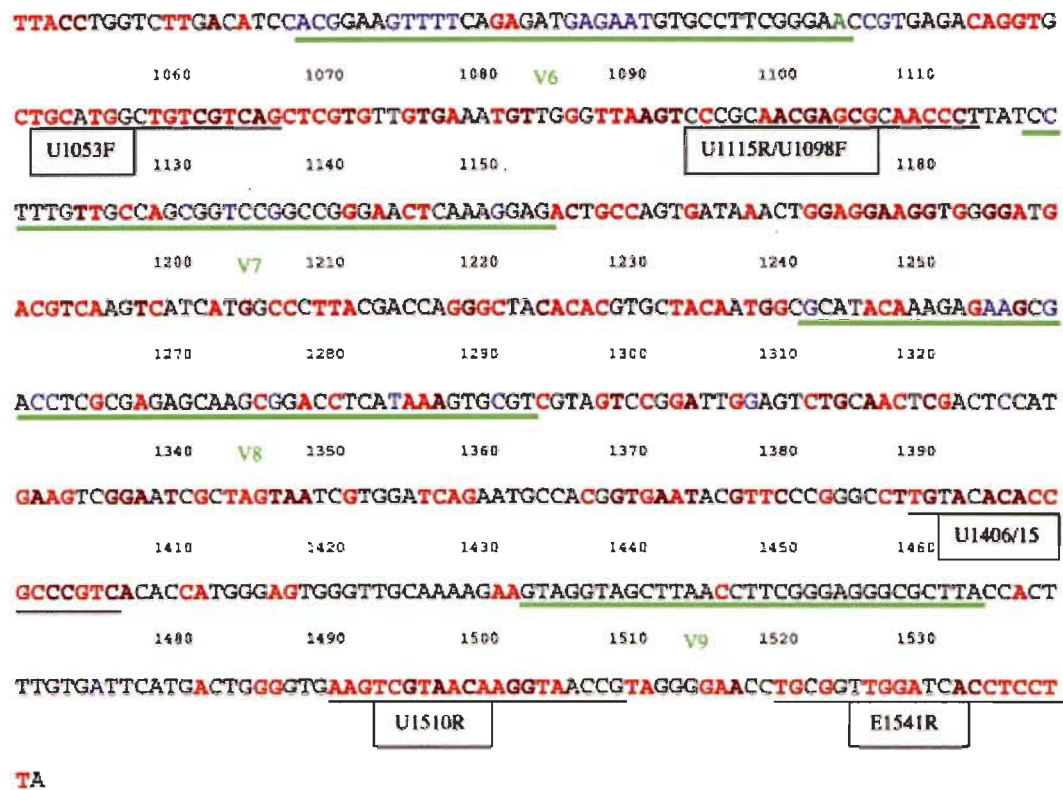


FIGURE 1.2 : Carte du gène ARNr16S de la bactérie *Escherichia coli* présentant les neuf régions hypervariables (soulignées en vert et numérotées de V1 à V9) [36].

1.5 MISE AU POINT D'UNE TRAPPE À PROMOTEURS BACTÉRIENS

Une technique mise au point dans notre laboratoire permet également la mise en évidence d'ADN bactérien à partir de prélèvements sanguins. Cette technique a déjà été utilisée pour évaluer l'expression de certains gènes [41], mais dans ce cas-ci, elle servira à mettre en évidence la présence de séquences promotrices bactériennes. Cette technique consiste à construire une trappe à promoteurs à l'aide d'un vecteur, pSP72-GFP (Green Fluorescent Protein) qui ne contient pas de promoteur en amont du gène codant pour la GFP. L'ADN génomique total extrait des échantillons sanguins est digéré avec une enzyme (*Sau3A* I), qui coupe à chaque séquence GATC, pour former des bouts compatibles aux bouts collants du plasmide digéré par l'enzyme de restriction *Bam*HI. Une ligase permet de lier les fragments d'ADN digérés et les plasmides pour ainsi

former des plasmides recombinants (vecteur pSP72-GFP + ADN génomique en amont du gène de la GFP). Théoriquement, si une séquence promotrice bactérienne est insérée en amont du gène codant pour la GFP, celle-ci sera produite et les bactéries transformées émettront de la fluorescence.

1.5.1 Promoteurs procaryotes versus eucaryotes

Pour que les bactéries émettent de la fluorescence, le plasmide recombinant doit avoir comme insert une séquence promotrice bactérienne capable d'activer la transcription du gène GFP. Théoriquement, seules les séquences procaryotes sont capables d'activer la transcription du gène puisque la machinerie cellulaire bactérienne est différente de celle des eucaryotes [42]. Premièrement, l'ARN polymérase devrait reconnaître les différentes séquences promotrices procaryotes (bactériens) et eucaryotes. Les promoteurs bactériens contiennent des motifs hexamériques aux positions -10 et -35 (par rapport au site d'initiation de la transcription) qui sont reconnus par l'ARN polymérase bactérienne [43]. La structure des promoteurs eucaryotes est beaucoup plus complexe et ils possèdent différents motifs : les boîtes TATA, INR, BRE, CCAAT et GC [42]. Pour que l'ARN polymérase bactérienne se lie au promoteur, les séquences promotrices doivent être d'origine bactérienne.

1.6 TRAVAUX ANTÉRIEURS PORTANT SUR LA DÉTECTION D'ADN BACTÉRIEN DANS LE SANG DE SUJETS SAINS

L'avènement de la biologie moléculaire a permis d'ouvrir plusieurs portes sur la détection de micro-organismes autrefois non identifiés par les méthodes conventionnelles. D'ailleurs, une étude effectuée par Nikkari et ses collègues, a permis de détecter de l'ADN bactérien dans le sang de quatre sujets sains. L'ADN génomique total du sang (récolté dans des tubes contenant un anti-coagulant) a été extrait et utilisé pour des réactions PCR avec des amorces ciblant le gène ADNr 16S. Les membres de sept groupes phylogénétiques et cinq divisions bactériennes ont été détectés dans le

sang. Les contrôles négatifs (eau) effectués ont aussi démontré une amplification lors de la réaction PCR. Par contre, certains phylotypes étaient présent seulement dans les échantillons sanguins : *Pseudomonas putida*, *Stenotrophomonas maltophila*, *Bacillus globisporus* et *Riemerella anatipestifer*. En définitive, les auteurs affirment qu'il est difficile de déterminer si l'ADN détecté provenait du fluorescente sang ou de la peau, ou des deux [44].

Lors d'une étude portant sur la recherche de spirochètes impliqués dans la maladie de l'Alzheimer, des chercheurs de l'Université McGill ont observé des bactéries pléomorphiques dans le sang de sujets sains. Le sang de 25 sujets sains a été prélevé dans des tubes ne contenant aucun anticoagulant. Les échantillons ont été observés au microscope à fond noir (Leitz Dialux 20) et l'ADN extrait pour des études moléculaires (PCR ciblant l'ADNr 16S et le gène *gyrB* et l'hybridation *in situ* (FISH)). Les résultats ont démontré la présence d'un phylotype apparenté à 99.6% à la bactérie *Stenotrophomonas maltophila* [45].

1.7 OBJECTIFS ET HYPOTHÈSES

L'omniprésence des bactéries n'est pas un fait nouveau, par contre, l'avènement de la biologie moléculaire a permis de caractériser plusieurs microflore que l'on ne connaissait pas auparavant [9, 12]. Les technologies moléculaires ont également mis en évidence un problème majeur relié aux transfusions sanguines : certains donneurs sains sont porteurs de bactéries transmises lors des transfusions [21, 24, 25]. De plus, deux études, mentionnées précédemment [44, 45] remettent en question la stérilité du sang. Ceci nous amène donc à se poser la question suivante: le sang humain est-il vraiment stérile? Puisque seulement deux études mettent en doute la stérilité du sang et que celles-ci n'ont pas démontré si l'ADN bactérien provenait du sang ou de la peau, une étude plus poussée permettrait d'élargir nos connaissances sur la possibilité que des bactéries vivent dans le sang d'humains en santé.

Notre hypothèse est donc la suivante : le sang d'un humain sain contient des bactéries vivantes. Le premier objectif est donc de mettre en évidence la présence d'ADN bactérien à l'aide de trois technologies : la PCR, la PCR quantitative et la trappe à promoteurs bactériens. Les deux premières technologies permettent de détecter l'ADN bactérien à l'aide d'amorces universelles ciblant le gène codant pour l'ARN ribosomal, plus particulièrement, la petite sous-unité 16S. Ces technologies ont déjà été utilisées et se sont avérées efficaces pour détecter de l'ADN bactérien dans le sang de sujets sains [44, 45]. Par contre, dans ce cas-ci, une autre technologie sera utilisée pour appuyer les résultats obtenus : la trappe à promoteurs bactériens. Un vecteur contenant le gène codant pour la GFP, sans son promoteur, permettra de mettre en évidence la présence de séquences promotrices bactériennes dans le sang.

La présence d'ADN bactérien ne démontre toutefois pas que les bactéries sont vivantes et originaires du sang. Ce qui nous amène au deuxième objectif : démontrer la viabilité des bactéries à l'aide de la PCR quantitative suite à une incubation. En effet, pour démontrer que l'ADN bactérien détecté n'est pas seulement un vestige de l'effet de la défense immunitaire, mais plutôt de l'ADN provenant de bactéries vivantes, une incubation suivie d'une détection d'ADN bactérien à l'aide de la PCR quantitative permettraient de démontrer qu'il y a une croissance bactérienne signifiant la présence de bactéries vivantes.

Puisque le but de cette étude est de démontrer la non-stérilité du sang, des mesures ont été prises tout au long des expériences pour éviter les contaminations extérieures et seront citées tout au long du texte.

CHAPITRE 2

MATÉRIEL ET MÉTHODES

2.1 LES CONTRÔLES

Les contrôles négatifs sont essentiels pour démontrer qu'il n'y a pas eu de contaminations lors des manipulations. Dans le cas présent, ces contrôles s'avèrent primordiaux puisque le but est de démontrer qu'il y a des bactéries dans le sang et que celles-ci ne sont pas présentes en raison de contaminations. Pour y parvenir, des contrôles négatifs, contenant de l'eau Gibco (EMD chemicals inc.) stérile autoclavée, ont été utilisés pour les différentes techniques : PCR conventionnelle, PCR quantitative et la trappe à promoteurs bactériens.

2.2 PRÉLÈVEMENTS DE SANG SUR DES SUJETS SAINS

Pour cette étude, le sang de 6 sujets, d'apparence saine (sans signe ou symptôme clinique suggérant une infection ou une maladie) et d'âge moyen (25-45 ans), a été prélevé au laboratoire d'hématologie de l'Université du Québec à Trois-Rivières. Le sang veineux a été prélevé, suite à une désinfection de la peau avec de l'alcool isopropylique à 70%, dans des tubes ne contenant aucun anti-coagulant (BD Vacutainer® Serum) et d'autres contenant les anti-coagulants suivant : EDTA (BD Vacutainer® K2 EDTA 7.2mg), héparine-sodium (BD Vacutainer® Sodium Heparin 68 USP units), citrate-sodium (BD Vacutainer® Buffered Cit. Na. (9:1)). Lors du prélèvement, deux tubes de sang ont été prélevés pour chaque individu et ce, pour les différents types de tubes (avec ou sans coagulant). Un tube a été utilisé au jour 0 (sans incubation) et l'autre a été incubé pour une période de 5 jours à une température de 30°C [45] dans le but de démontrer la prolifération bactérienne. De plus, l'avantage d'avoir deux tubes est de diminuer les risques de contaminations lors des manipulations : le tube incubé n'est pas ouvert jusqu'au moment de l'extraction. Toutes les manipulations du sang se font sous

une hotte pour empêcher la contamination extérieure. Comme cité précédemment, des contrôles négatifs ont été effectués en ajoutant de l'eau Gibco (EMD Chemicals inc.) stérile autoclavée aux différents tubes de prélèvements et ceux-ci ont subi les mêmes traitements que ceux contenant le sang.

2.3 EXTRACTION DE L'ADN DU SANG

Une trousse d'extraction d'ADN génomique du sang (QIAamp[®] DNA Blood Mini Kit, QIAGEN) a été utilisée. Un volume de 20µl de protéase QIAGEN[®] a été pipeté dans un microtube de 1,5ml. Cette enzyme permet de scinder les protéines et elle conserve son activité en présence de différents agents comme l'EDTA. Ensuite, un volume de 200µl de sang (provenant des différents prélèvements) a été ajouté au microtube. Le tampon AL (200µl) permettant la lyse des cellules a été ajouté à l'échantillon pour ensuite être vortexé pour une période de 15 secondes. Le tout a été incubé 10 minutes à 56°C et, par la suite, centrifugé brièvement. De l'éthanol 100% (200µl) a été ajouté au microtube et vortexé pendant 15 secondes.

À la suite d'une brève centrifugation, l'échantillon a été transféré dans une colonne QIAamp et centrifugé à 8000rpm pour 1 minute. La colonne permet de recueillir l'ADN sur la membrane de gel de silice et ce, grâce à un gradient de salinité et de pH. Ainsi, seulement l'ADN adhère à la membrane donc les protéines et autres contaminants, pouvant nuire à la réaction PCR, ne sont pas retenus.

Ensuite, plusieurs lavages ont été effectués avec différents tampons pour purifier l'ADN présent sur la membrane. Premièrement, suite à une première centrifugation, le tampon AW1 (500µl) a été ajouté à la colonne et centrifugé à 8000rpm pour 1 minute. Ensuite, la colonne a été transférée dans un nouveau tube collecteur (pour éviter l'accumulation d'éthanol pouvant nuire à la réaction PCR). Un volume de 500µl de tampon AW2 a été ajouté à la colonne et centrifugé à 14 000rpm pour 3 minutes. Une autre centrifugation a été effectuée à la même vitesse pour 1 minute pour s'assurer qu'il n'y ait pas de résidus.

Pour éluer l'ADN, la colonne a été transférée dans un autre microtube collecteur et 100µl de tampon AE a été ajouté. Une incubation de 5 minutes à température de la pièce a permis de bien concentrer l'ADN pour ensuite, le recueillir après une centrifugation à 8000rpm pour 1 minute. L'éluant contenant l'ADN génomique total du sang a été récolté et utilisé pour les étapes suivantes. L'ADN non utilisé a été conservé à 4°C. Les différentes étapes sont présentées à la figure 2.1.

2.3.1 Efficacité de l'extraction

Pour vérifier l'efficacité de la trousse d'extraction, des contrôles ont été effectués avec des amorces ciblant le gène de la β -actine (tableau 2.1). Une réaction PCR (réactifs présentés au tableau 2.2 et paramètres au tableau 2.3) avec de l'ADN génomique, extrait du sang au jour 0 et au jour 5, a été effectuée. Des amorces, ciblant une région de 348 nucléotides située sur le gène codant pour la β -actine (présent chez les eucaryotes), ont été utilisées. Un contrôle positif a été fait avec de l'ADN eucaryote complémentaire (ADNc 370) et un contrôle négatif avec de l'eau Gibco stérile.

2.3.1.1 Migration des produits PCR et photographie

Suite à la réaction PCR, 5µl de chaque produit PCR (échantillons et contrôles) ont été ajoutés à 2µl de bleu de migration (bleu de bromophénol 0.25% p/v, xylène cyanol FF 0.25% p/v, glycérol 50% p/v et EDTA 0.2M). Pour le marqueur (« DNA ladder »), dans ce cas-ci 100pb (500µg/ml, NEB), seulement 1µl a été ajouté au bleu de migration (2µl). Les puits du gel d'agarose LE 0.8% (Roche), dilué dans du tampon TBE 1X, ont été remplis avec les différents échantillons, les contrôles et le marqueur. La migration a été effectuée dans un bac à migration contenant du TBE 1X (Tris 1M, acide borique 1M et EDTA 0.2M). Le voltage a été réglé à 120 volts pour une durée de 60 minutes.

The QIAamp DNA Blood Mini Spin and Vacuum Procedures

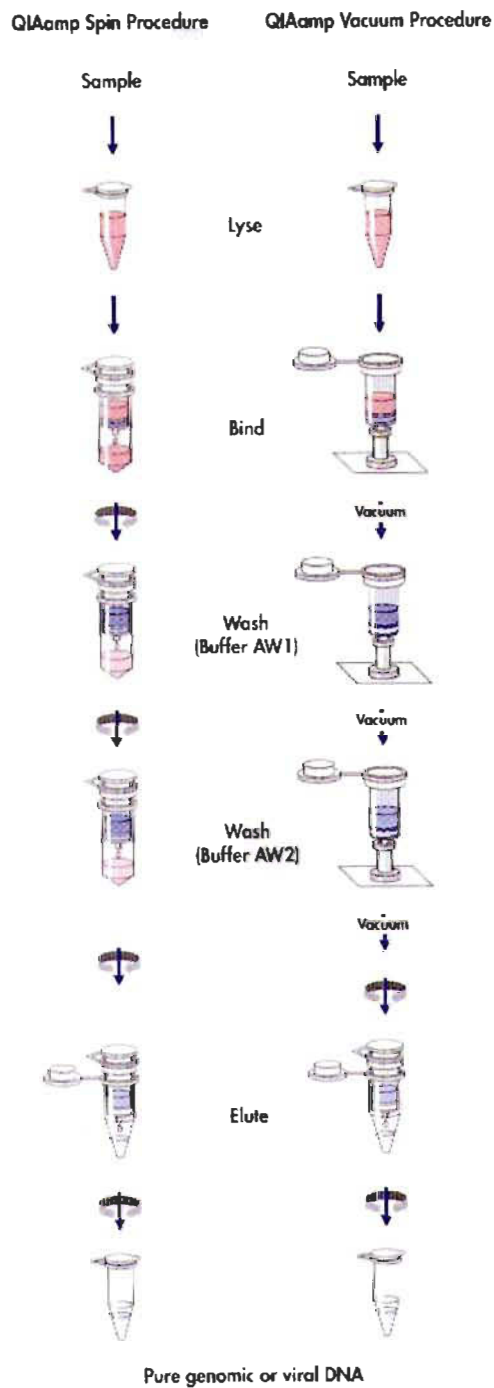


FIGURE 2.1 : Schéma illustrant les étapes de l'extraction d'ADN génomique du sang avec la trousse QIAamp DNA Blood Mini Kit.

TABLEAU 2.1
Caractéristiques des amorces b-actine

Amorces	Séquence	%GC
b-actine sens	5'-GAG GAT CTT CAT GAG GTA GTC TGT CAG GTC - 3'	48%
b-actine anti-sens	5'-CAA CTG GGA CGA CAT GGA GAA GAT CTG GCA- 3'	52%

TABLEAU 2.2
Réactifs utilisés lors de la réaction PCR avec les amorces b-actine

Réactifs	Volumes (µl) (échantillons et contrôle positif)	Volumes (µl) (contrôle négatif)
Eau Gibco stérile (EMD)	33.5	38.5
Tampon ThermoPol 10X (NEB)	5	5
dNTP 10mM (NEB)	1	1
Amorces b-actine sens 15µM (NEB)	2.5	2.5
Amorces b-actine anti-sens 15µM (NEB)	2.5	2.5
ADN génomique ou complémentaire	5	-----
<i>Taq</i> DNA polymerase 5000U/ml (NEB)	0.5	0.5
Total	50	50

TABLEAU 2.3
Paramètres de l'appareil PTC100 (MJ researchs)

Étape	Température (°C)	Durée	
dénaturation	94	3 minutes	
	94	45 secondes	
Appariement des amorces	58	30 secondes	30 cycles
élongation	72	1 minute	
	72	4 minutes	

Suite à la migration, le gel a été transféré dans un bac contenant 0.01% de colorant à gel SYBR[®]Safe (10 000X concentré dans du sulfoside de diméthyl (DMSO), Invitrogen) dilué dans du tampon TBE 1X. Le tout a été agité sur une plaque agitatrice pour une durée de 20 minutes. Ensuite, le gel a été photographié à l'aide de l'appareil d'imagerie moléculaire (Molecular Imager Gel DOC system, BIORAD). Le logiciel « quantity one » a permis d'ajuster les paramètres de la caméra.

2.4 PCR CONVENTIONNELLE

Pour détecter de l'ADN bactérien dans le sang, l'ADN génomique total (extrait du sang des différents sujets à partir de tubes sans coagulant (BD Vacutainer[®] Serum) et avec coagulant (BD Vacutainer[®] K2 EDTA 7.2mg) au jour 0 et 5) a été utilisé. Plusieurs réactions PCR ont été effectuées avec des amorces ciblant le gène ARNr16S et amplifiant une région de 1505 nucléotides qui comprend les 9 régions variables. Les amorces sont décrites dans le tableau 2.4.

TABLEAU 2.4
Caractéristiques des amorces utilisées pour le PCR conventionnel

Amorces	Séquence	%GC
For 16S RNA <i>E. coli</i>	5'-AGA GTT TGA TCC TGG CTC AG- 3'	50%
Rev 16S RNA <i>E. coli</i>	5'-AAG GAG GTG ATC CAG CCG CA- 3'	60%

De nombreuses réactions PCR ont été effectuées avec différents paramètres pour assurer une amplification plus efficace. Premièrement, les amorces ont été testées sur l'ADN génomique de différentes bactéries ((a) *Actinobacillus pleuropneumoniae* 4074 (*App*4074), (b) *E. coli* EC100, (c) *E. coli* DH5 α et (d) des souches bactériennes non identifiées) pour mettre au point les paramètres. L'optimisation de la réaction PCR constitue un facteur important pour obtenir une bonne amplification. La composition des amorces et la longueur de l'amplicon (produit PCR amplifié) contribuent au choix des différents paramètres de la réaction PCR. La température d'appariement (T_A) dépend de

la composition en GC des amorces. Le calcul suivant permet de calculer la T_M («melting temperature») : $2 (A+T) + 4 (G+C)$. Pour les amorces «For 16S RNA *E. coli*» et «Rev 16S RNA *E. coli*», la T_M est de 60°C et 64°C, respectivement. La différence de température entre les 2 amorces ne doit pas différer de plus de 5°C. En résumé, la T_M est de 60°C puisque l'on se réfère à la plus basse des deux. La T_A se situe habituellement à 2°C à 10°C sous la T_M . Pour cette réaction, la température d'appariement se situerait entre 50°C et 58°C. Les premières réactions avec les contrôles positifs (ADNs génomiques bactériens) ont été effectuées à une T_A de 50°C.

Les différents volumes des réactifs utilisés lors des PCRs sont présentés au tableau 2.5 et les différents paramètres de l'appareil figurent au tableau 2.6. Un contrôle négatif a été effectué avec de l'eau Gibco stérile autoclavée et un contrôle positif avec de l'ADN génomique (2 ou 5µl) provenant de la bactérie *E. coli* DH5α (≈ 100ng/µl). Pour la migration et la photographie des gels, se référer au paragraphe 2.3.1.1. Le marqueur 1KB (500µg/ml, NEB) a été utilisé. Les produits PCR (5µl et 10µl), effectués à partir des échantillons, ont été ajoutés aux puits du gel.

2.4.1 Purification des produits PCR

Pour identifier la séquence amplifiée, les produits PCR (≈ 1.5kb) ont été purifiés pour permettre leur insertion dans le vecteur pDrive (QIAGEN®). Les vecteurs recombinant contenant les produits PCR ont ensuite été envoyés au séquençage. Pour purifier l'ADN provenant des produits PCR, une trousse de purification (QIAquick® PCR Purification Kit, QIAGEN®) a été utilisée.

Premièrement, 5 volumes de tampon PBI ont été ajoutés à 1 volume de produit PCR. L'échantillon a été transféré dans une colonne QIAquick et centrifugé à une vitesse de 13000rpm pendant 30 secondes. L'ADN provenant du produit PCR a ainsi été fixé à la membrane de silice et les produits résiduels éliminés (amorces, nucléotides, polymérases et sels) lors de la centrifugation.

Pour laver l'ADN, un volume de 750µl du tampon PE a été ajouté au tube contenant la colonne et centrifugé 60 secondes à 13000rpm. Le liquide résiduel a été jeté et une autre centrifugation a été effectuée pour s'assurer qu'il n'y ait aucun résidu.

L'ADN a été élué dans un nouveau microtube avec 30µl de tampon EB. Le tout a été incubé 1 minute à température de la pièce avant d'être centrifugé 1 minute à 13000rpm. Les produits PCR purifiés ont été utilisés pour être cloné dans le vecteur pDRIVE (QIAGEN®).

TABLEAU 2.5
Réactifs utilisés lors des réactions PCR avec les amorces 16S RNA E. coli

Réactifs	Réaction 1		Réaction 2	
	Volumes (µl) Échantillons	Volumes (µl) Contrôle -	Volumes (µl) échantillons	Volumes (µl) Contrôle -
Eau Gibco stérile (EMD)	39.5	41.5	33.5	38.5
Tampon ThermoPol 10X (NEB)	5	5	5	5
dNTP 10mM (NEB)	1	1	1	1
Amorces For 16S RNA E.coli 15µM (NEB)	1	1	2.5	2.5
Amorces Rev 16S RNA E. coli 15µM (NEB)	1	1	2.5	2.5
ADN génomique extrait du sang ¹	2	-----	5	-----
Taq DNA polymerase 5000U/ml (NEB)	0.5	0.5	0.5	0.5
Total	50	50	50	50

1. La concentration d'ADN génomique pour les échantillons provenant du sang est $\approx 20\text{ng}/\mu\text{l}$ tandis que la concentration d'ADN génomique provenant de bactéries est $\approx 100\text{ng}/\mu\text{l}$

TABLEAU 2.6
Différents paramètres utilisés lors des réactions PCR
effectuées avec le PTC100 (MJ researchs)

Étape ¹	Paramètres 1		Paramètres 2		Paramètres 3		Paramètres 4	
	T (°C)	Durée	T (°C)	Durée	T (°C)	Durée	T (°C)	Durée
Dénaturation	94	3 min	94	3 min.	94	3 min.	94	3 min.
	94	30 sec	94	30sec.	94	30 sec.	94	45 sec.
Appariement	50	30 sec	48	30 sec	45	30 sec.	45	30 sec.
Élongation	72	2 min.	72	2 min.	72	2 min.	72	2 min.
	72	7 min.	72	7 min.	72	7 min.	72	7 min.

1. le nombre de cycles effectués est de 40 pour les 4 réactions PCR

2.4.2 Clonage dans le vecteur pDrive

Les produits PCR purifiés ont été insérés dans le vecteur pDRIVE (figure 2.2) en effectuant une ligation et ce, à l'aide de la trousse de clonage QIAGEN[®] PCR Cloning^{plus} Kit (QIAGEN[®]). Le vecteur est idéal pour les produits PCR puisque celui-ci contient un uracile (U) à chaque bout pour permettre l'insertion du produit PCR poly-adenylé (AAA) par la *Taq* polymérase. De plus, il contient un gène de résistance à l'ampicilline permettant sa sélection sur un milieu LB (Luria Bertani) sélectif contenant de l'ampicilline (100µg/ml) et le gène *lacZ* permet de vérifier si le produit PCR a été inséré : s'il n'est pas inséré, la β -galactodidase sera exprimée et les colonies formées sur un milieu contenant de l'IPTG (isopropyl β -D-thiogalactopyranoside) à 50µM et du X-Gal (5-bromo-4-chloro-3-indolyl β -D-thiogalactopyranoside) à 80µg/ml seront bleues, tandis que s'il est inséré, la β -galactodidase ne sera pas produite et les colonies seront blanches.

Pour effectuer la ligation, les réactifs au tableau 2.7 ont été utilisés. Le tout a été incubé à 16°C pour 2 heures. Pour vérifier si le produit PCR a bien été inséré, la ligation a été électroporée à l'intérieur de cellules bactériennes compétentes QIAGEN EZ.

TABLEAU 2.7
Réactifs utilisés et leur volume pour le clonage
des produits PCR dans le vecteur pDRIVE

Réactifs	Volume (µl)
Vecteur pDRIVE (50ng/µl)	1
Produits PCR purifiés (30ng/µl)	4
Mélange pour ligation (enzyme + tampon) « Ligation Master Mix 2X »	5
Total	10

2.4.2.1 Préparation des cellules bactériennes compétentes QIAGEN EZ

Pour préparer les cellules compétentes, 2 à 3 colonies de cellules QIAGEN EZ, provenant d'une gélose TSA (Tryptic Soy Agar, BD) avec du sang de mouton défibrinisé (5%), ont étéensemencées dans un erlenmayer de 500ml contenant 4ml de bouillon TSB (Tryptic Soy Broth, BD). Le tout a été incubé 16 heures à 37°C.

Ensuite, 36ml de TSB ont été ajoutés à l'erlenmayer pour totaliser un volume de 40ml qui a été incubé 90 minutes à 37°C. La culture a été transférée dans un tube de 40ml et gardé sur glace. Une première centrifugation à 4°C (vitesse de 3000rpm pour 20 minutes) a été effectuée pour récolter les cellules en décantant le surnageant et gardant le culot intact. Celui-ci a été resuspendu dans un volume d'eau stérile à 4°C pour laver les cellules. Une autre centrifugation a été effectuée aux mêmes paramètres que précédemment. Un autre lavage a été effectué et cette fois, le culot résiduel a été resuspendu dans le volume d'eau résiduel. Un volume de 40µl de cellules compétentes a été utilisé pour chaque électroporation.

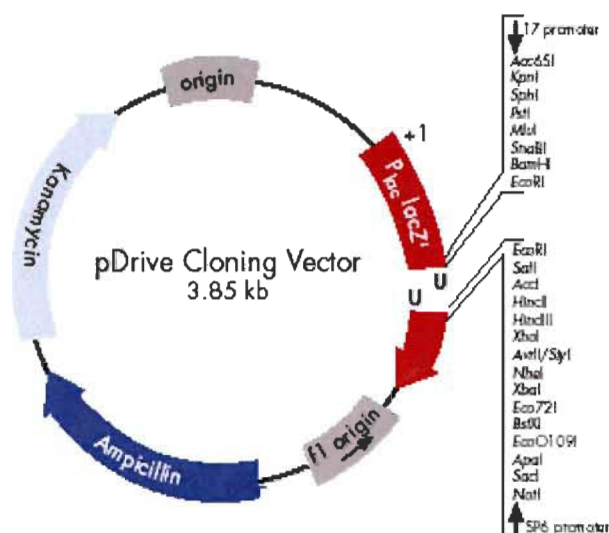


FIGURE 2.2 : Vecteur de clonage pDRIVE (source : <http://www1.qiagen.com/>).

2.4.2.2 Électroporation des ligations dans les cellules compétentes QIAGEN EZ

Un volume de 2.5µl de chaque ligation, effectuée sur les différents produits PCR, a été ajouté au 40µl de cellules QIAGEN EZ. Pour vérifier la qualité des géloses LB ampicilline (100µg/ml), X-Gal (80µg/ml) et IPTG (50µM), 1µl du vecteur pUC18 (100ng/µl) a été ajouté à un autre tube contenant aussi 40µl de cellules QIAGEN EZ. Puisque ce vecteur contient également le gène codant pour la β -galactodidase, les bactéries l'ayant incorporé forment des colonies bleues sur ce milieu. De plus, un contrôle négatif (contenant seulement des cellules QIAGEN EZ) a été effectué pour s'assurer que les cellules compétentes n'aient pas été contaminées. Aucune colonie ne devrait croître puisque les cellules ne contiennent pas le vecteur permettant la croissance sur un milieu contenant de l'ampicilline.

Pour l'électroporation, les échantillons contenant les ligations et les contrôles ont été transférés dans des cuvettes à électroporation de 1mm (VWR Scientifics Products). L'appareil *E.coli* pulser (BIO RAD) a été réglé à un voltage de 1.8kv. Immédiatement après l'envoi du courant, 1ml de TSB a été ajouté pour permettre aux cellules de

survivre. Ensuite, les cellules ont été transférées dans un microtube pour ensuite être incubées à 37°C pour une période de 90 minutes. À la suite de cette incubation, 100µl de chaque échantillon et contrôle ont été étalés sur un milieu LB ampicilline (100µg/ml), X-Gal (80µg/ml) et IPTG (50µM). Une autre incubation de 16 heures à 37°C a été effectuée pour permettre aux bactéries de croître et ainsi déterminer la présence de colonies blanches contenant le vecteur ayant incorporé le produit PCR.

2.4.2.3 « Cracking » pour confirmer l'insertion du produit PCR dans pDRIVE

L'avantage du « cracking » est qu'il permet d'évaluer rapidement la taille des plasmides (vecteurs). Avant de commencer ce protocole, il est important de faire des patchs à partir des colonies sélectionnées, dans ce cas-ci, les colonies blanches. Pour ce faire, une colonie est prélevée et ensemencée sur une superficie de 1cm² sur le même milieu (LB ampicilline (100µg/ml), X-Gal (80µg/ml) et IPTG (50µM)) et ce, à l'aide d'un cure-dent. Puisque le « cracking » nécessite une grande quantité de bactéries, ceci permet d'augmenter la population bactérienne d'une colonie en particulier. Une partie de la patch est prélevée et resuspendue dans 25µl d'eau stérile. La solution de « cracking » est ensuite préparée en ajoutant 9 volumes de tampon de cracking (EDTA 0.5M pH8, SDS 0.04M, bleu de bromophénol 10%, glycérol 10%) à 1 volume de NaOH 1N. Un volume de 25µl est ensuite ajouté à la suspension bactérienne et le tout bien mélangé. Ensuite, 25µl est déposé dans des puits sur gel d'agarose (0.08%). Comme marqueur de poids moléculaire, 2µl du supercoiled (0.25µg/µl, Invitrogen) a été utilisé. La migration s'est effectuée à un voltage de 120 volts pour une période de 60 minutes. Pour la coloration et les photographies du gel, se référer au paragraphe 2.3.1.1. Puisque le vecteur a une taille de 3.85kb, la taille du vecteur contenant le produit PCR est d'approximativement 5.35kb.

2.4.2.4 Mini-prep pour purifier l'ADN plasmidique (vecteur pDRIVE)

Pour permettre un bon séquençage et ainsi identifier l'origine de la séquence insérée (produit PCR), les vecteurs contenant le produit PCR ont été purifiés à l'aide d'une trousse de purification d'ADN : QIAprep[®] Miniprep, QIAGEN[®].

Pour commencer, une partie de la patch (utilisée précédemment pour le « cracking ») a été prélevée et ensemencée dans 10ml de milieu LB ampicilline (100µg/ml) pour sélectionner seulement les vecteurs contenant le gène de résistance à l'ampicilline. La culture a ensuite incubée 16 heures à 37°C avec agitation (120rpm). Pour récolter les cellules bactériennes, les tubes, contenant le 10ml de culture, ont été centrifugés à une vitesse de 3500rpm pour 20 minutes. Le liquide a été jeté pour laisser seulement le culot contenant les cellules bactériennes.

Le culot a été resuspendu dans 500µl de tampon P1 et transféré dans un microtube. Un volume de 500µl de tampon P2 a été ajouté et mélangé en inversant doucement le tube 6 fois. Ces derniers tampons permettent une lyse bactérienne alcaline libérant ainsi l'ADN. Le tout a été neutralisé avec l'ajout du tampon N3 (700µl). Immédiatement après l'ajout du neutralisant, la solution a été mélangée en inversant le tube 6 fois de plus. La solution, devenue embrouillée, a été centrifugée 10 minutes à 13 000rpm.

Le surnageant contenant l'ADN plasmidique (vecteur) a ensuite été transféré dans une colonne QIAprep qui contient une membrane de silice capable de retenir l'ADN. Suite à une centrifugation (60 secondes à 13 000rpm), la colonne a été lavée avec 750µl de tampon PE et centrifugée à nouveau. Une centrifugation supplémentaire a été effectuée pour s'assurer qu'il n'y ait plus de tampon résiduel.

Pour éluer l'ADN plasmidique, la colonne a été transférée dans un nouveau microtube et 30µl de tampon EB y ont été ajoutés. Pour recueillir l'éluant, une centrifugation à 13 000rpm pour 60 secondes a été effectuée. L'ADN plasmidique a

ensuite été dosé au spectrophotomètre (GeneQuant pro) à une longueur d'onde de 260nm.

2.4.2.5 Préparation des échantillons pour le séquençage

Une fois les vecteurs purifiés, ceux-ci ont été dosés et parfois dilués dans du tampon EB pour atteindre une concentration de 100ng/μl, idéale pour le séquençage. Les amorces ont également été dosées pour obtenir une concentration de 1.5μM. Dans ce cas-ci, les amorces utilisées ont été : « For 16S RNA E.coli » et « Rev 16S RNA E. coli ». Pour chaque réaction de séquençage (sens et anti-sens), 10μl d'ADN plasmidique ont été utilisés et 5μl de chaque amorce. Les échantillons préparés ont été envoyés à la plate-forme d'analyses biomoléculaires de l'Université Laval (www.bioinfo.ulaval.ca). Pour permettre l'identification du produit PCR introduit dans le vecteur, la séquence analysée a été comparée (*BLAST* = Basic Local Alignment Search Tool) aux séquences génomiques humaines et bactériennes présentes dans la banque de données NCBI (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>)

2.5 PCR QUANTITATIVE

Pour détecter la présence de bactéries vivantes, l'ADN génomique total (10 ng/μl) extrait du sang (sans coagulant ou avec différents coagulants : héparine-sodium (BD Vacutainer® Sodium Heparin 68 USP units) et citrate-sodium (BD Vacutainer® Buffered Cit. Na. (9:1))) des sujets (jour 0 et jour 5) a d'abord été amplifié par la PCR quantitative. Les amorces ciblent le gène codant pour l'ARNr 16S et amplifient une région de 367 nucléotides qui comprend les régions variables V7 à V9. Pour cette réaction, une concentration de 0.7μM pour chaque amorce a été utilisée.

TABLEAU 2.8
Caractéristiques des amorces pour la PCR quantitative

Amorces	Séquence	%GC
p1 lpc	5'-GAG GAA GGT GGG GAT GAC GT- 3'	60%
Rev 16S RNA E.coli	5'-AAG GAG GTG ATC CAG CCG CA- 3'	60%

Pour préparer les échantillons, les réactifs suivants ont été utilisés (tableau 2.9 pour les quantités) : une solution de SYBR®Green (QuantiTect™ SYBR®Green PCR kit, Qiagen) contenant l'enzyme (HotStarTaq®DNA polymerase), le tampon (QuantiTect SYBR Green PCR Buffer), les colorants fluorescents (SYBR Green I et ROX), les dinucléotides triphosphate (dNTP) et le MgCl₂ (dichlorure de magnésium). La réaction a été effectuée avec l'appareil Lightcycler (Roche) avec les paramètres suivants : une première dénaturation à 95°C pour 15 minutes et 50 cycles comprenant; une seconde dénaturation à 94°C (15 secondes), l'appariement des amorces à 55°C (20 secondes) et une élongation à 72°C (35 secondes).

TABLEAU 2.9
Différents réactifs et volumes utilisés pour la PCR quantitative

Réactifs	Volume (µl)
2X QuantiTect SYBRGreen PCR	10
Amorce p1 lpc 0.7µM (NEB)	0.77
Amorce Rev 16S RNA E.coli 0.7µM (NEB)	0.77
Eau Gibco stérile (EMD)	6.46
Total	18

Le volume total par capillaire est de 20µl, donc il reste 2µl pour l'ADN. Dans ce cas ci, l'ADN extrait du sang (10ng/µl) a été ajouté pour les différents échantillons à tester. Un contrôle positif avec de l'ADN génomique (2µl) provenant de la bactérie *E. coli* DH5a (0.05pg/µl) a été effectué. Comme contrôle négatif, de l'eau Gibco stérile autoclavée (2µl), ayant subi le même protocole d'extraction que l'ADN provenant du

sang, a été utilisée pour vérifier s'il y a eu une contamination des réactifs lors des manipulations. Les résultats ont été analysés par le logiciel « *Roche Molecular Biochemicals Lightcycler Software* » sous forme de graphique logarithmique représentant la fluorescence émise en fonction du nombre de cycles effectués.

2.6 TRAPPE À PROMOTEURS BACTÉRIENS

La seconde technologie permet également de mettre en évidence de l'ADN bactérien dans le sang, mais ce, de manière aléatoire. Il s'agit de la trappe à promoteurs bactériens : le vecteur pSP72-GFP (créé dans notre laboratoire) qui ne contient pas de promoteur pour le gène codant pour la GFP, a été utilisé.

2.6.1 Construction du vecteur pSP72-GFP

Le vecteur pSP72-GFP a été construit dans notre laboratoire et il est une combinaison du vecteur pSP72 (figure 2.3) et du vecteur pGreenTIR (figure 2.5). Le gène codant pour la GFP (751 paires de base (pb)) a été enlevé du vecteur pGreenTIR en digérant celui-ci avec l'enzyme de restriction *EcoRI* puisque ses sites se situent exactement avant et après le gène d'intérêt ainsi, le promoteur (Plac) n'a pas été inclus. Ensuite, la cassette GFP, contenant la région initiale de traduction (TIR) et la séquence de Shine-Dalgarno (SD), a été introduite dans le site de polyclonage, plus précisément à la position 67 (*EcoRI*) du vecteur pSP72 et ce, à l'aide d'une ligase. La figure 2.7 présente le vecteur recombinant qui contient un gène sélectif pour l'ampicilline, permettant ainsi la sélection des bactéries contenant le vecteur suite à une transformation par électroporation. Le vecteur contient aussi le gène de la GFP et en amont, il y a un site de polyclonage pour permettre l'insertion de séquences.

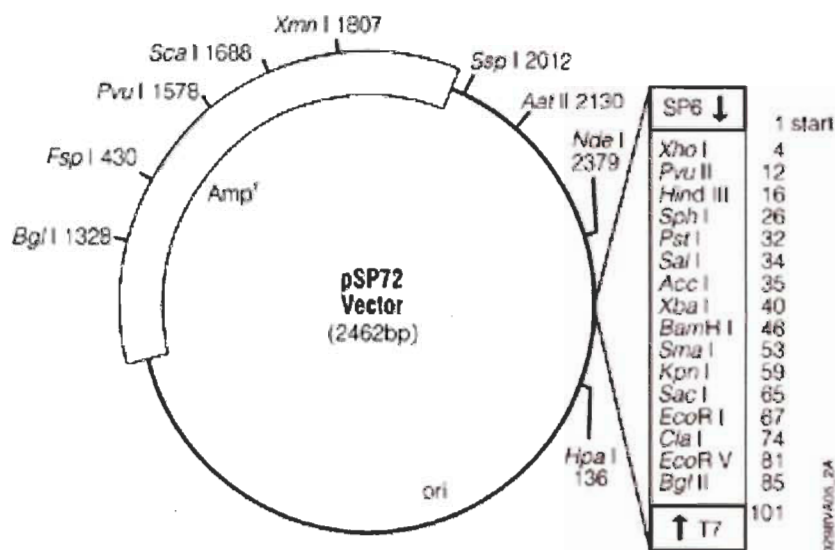


FIGURE 2.3 : Carte du vecteur pSP72 (<http://www.promega.com/tbs/tb040/tb040.pdf>).

2.6.1.1 Digestion et ligation

Une digestion enzymatique BamH1 (figure 2.4) (5U/réaction, NEB) a été effectuée sur le vecteur (1500ng utilisé). Pour ce faire, le vecteur pSP72-GFP, le tampon NEB buffer 3 (1X, NEB), le tampon BSA (10X, NEB) et l'enzyme BamH1 ont été mélangés dans un microtube et incubés 2 heures à 37°C. Pour désactiver l'enzyme, une purification PCR a été effectuée à l'aide de la trousse QIAquick®PCR purification (QIAGEN®). Premièrement, 5 volumes de tampon PBI ont été ajoutés à 1 volume de produit de digestion. Le tout a été transféré dans une colonne QIAquick et centrifugé à 13 000rpm pour 1 minute. Le liquide a été jeté et un volume de 750µl de tampon PE a été ajouté pour laver le produit de digestion. Une centrifugation à 13 000rpm pour 1 minute a été effectuée. Pour éluer, 30µl de tampon EB ont été ajoutés à la colonne et le tout a reposé 1 minute à température de la pièce. Le produit de digestion a été récupéré en centrifugeant 1 minute à la même vitesse que précédemment.



FIGURE 2.4 : Sites de reconnaissance de l'enzyme BamHI.

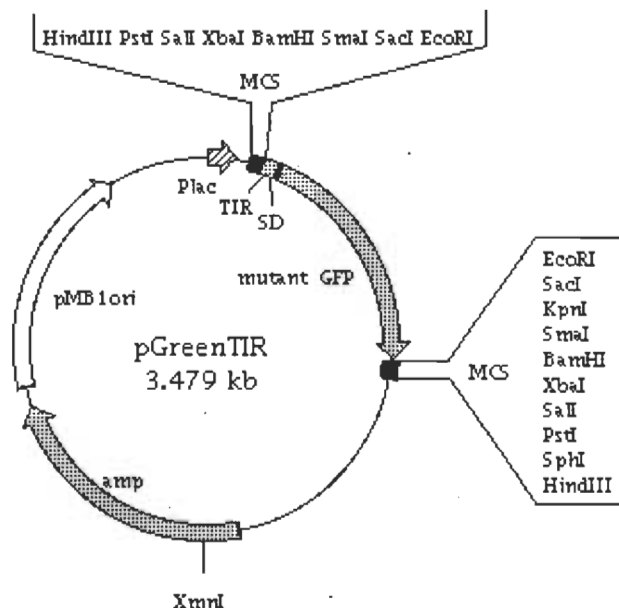


FIGURE 2.5 : Carte du vecteur pGreenTIR (<http://gillnet.lab.nig.ac.jp/~cvector/map/pGreenTIR.gif>).

Le vecteur pSP72-GFP digéré (1500ng utilisé) a été déphosphorylé à l'aide de la phosphatase alcaline (10U/réaction, NEB) et du tampon phosphatase 10X pour empêcher le vecteur de se refermer sur lui-même. Une incubation à 37°C pour 15 minutes a permis à la phosphatase d'agir et une désactivation à 65°C pour 15 minutes a été effectuée.

L'ADN génomique, provenant des extractions effectuées à partir du sang (avec différents coagulants : EDTA (BD Vacutainer® K2 EDTA 7.2mg), héparine-sodium (BD Vacutainer® Sodium Heparin 68 USP units), citrate-sodium (BD Vacutainer® Buffered Cit. Na. (9:1))) de 4 sujets au jour 0, a été digéré ($\approx 350\text{ng/}$ réaction) avec l'enzyme Sau3AI (figure 2.6) (10U/réaction, NEB). Les digestions se sont effectuées en ajoutant l'ADN extrait, le tampon Sau3AI 1X (NEB), le tampon BSA 10X (NEB) et l'enzyme

dans un microtube et le tout a été incubé 2 heures à 37°C. Pour désactiver, une incubation à 65°C a été effectuée.



FIGURE 2.6 : Sites de reconnaissance de l'enzyme Sau3AI.

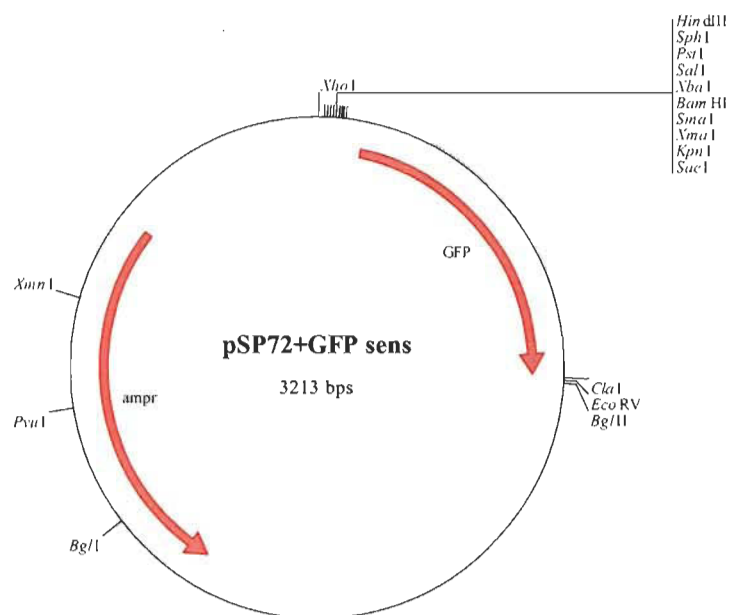


FIGURE 2.7 : Carte du vecteur pSP72-GFP (construit en laboratoire).

Une ligation, avec la T4 ligase (5U/réaction, NEB), le tampon de ligase T4 1X (NEB), a permis de créer une banque génomique contenant des vecteurs pSP72-GFP (100ng utilisé) recombinants : l'ADN digéré (80ng utilisé) a généré plusieurs fragments qui ont été insérés en amont du gène, dans le site de polyclonage. La ligation s'est effectuée à 16°C pendant 16 heures et l'enzyme a été désactivée à 65°C pendant 20 minutes.

2.6.1.2 Électroporation des ligations dans la bactérie *E. coli* DH5α

Les ligations (2.5µl) ont ensuite été électroporées (paragraphe 2.4.2.2) dans des cellules bactériennes électrocompétentes, dans ce cas-ci, *E. coli* DH5α. Le protocole pour la préparation des cellules est le même qu'au paragraphe 2.4.2.1. À la suite de l'électroporation, il y a eu une incubation de 30 minutes (au lieu de 60 minutes). Un contrôle négatif a également été effectué pour vérifier la présence de contaminations. Le plasmide pUC18 a été utilisé comme contrôle pour vérifier la qualité des milieux utilisés : LB ampicilline (100ng/µl). Un volume de 100µl de chaque ligation (8 par ligation) et de chaque contrôle a étéensemencé sur le milieu sélectif et une incubation de 16 heures à 37°C a été effectuée.

Les colonies fluorescentes aux Ultra-Violets (UVs) ont été sélectionnées pour faire des patchs dans le but d'évaluer la taille des vecteurs recombinants par « cracking » (paragraphe 2.4.2.3). Les vecteurs contenant un fragment d'ADN digéré ont été purifiés par une mini-prep (paragraphe 2.4.2.4). Les vecteurs recombinants purifiés ont ensuite été séquencés à l'aide de deux amorces (tableau 2.9) qui ciblent une région en aval du site de restriction BamH1 (Rev pSP72-GFP) et une en amont du site de restriction (For pSP72-GFP). La préparation pour le séquençage est mentionnée au paragraphe 2.4.2.5 et les échantillons ont été envoyés à la plate-forme d'analyses biomoléculaires de l'université Laval.

TABLEAU 2.10
Caractéristiques des amorces utilisées pour le séquençage
des fragments d'ADN insérés dans le vecteur pSP72-GFP

Amorces	Séquence	%GC
For pSP72-GFP	5'-TTG TCG TTA GAA CGC GGC TA- 3'	50%
Rev pSP72-GFP	5'-GGG ACA ACT CCA GTG AAA AG- 3'	50%

CHAPITRE 3

PCR CONVENTIONNELLE POUR LA DÉTECTION D'ADN BACTÉRIEN

3.1 VÉRIFICATION DE L'EFFICACITÉ DE L'EXTRACTION D'ADN GÉNOMIQUE DU SANG

Pour s'assurer de l'efficacité de l'extraction d'ADN génomique, effectuée à partir des échantillons de sang veineux, un contrôle a été effectué en exécutant une PCR avec des amorces ciblant le gène codant pour la β -actine (figure 3.1).

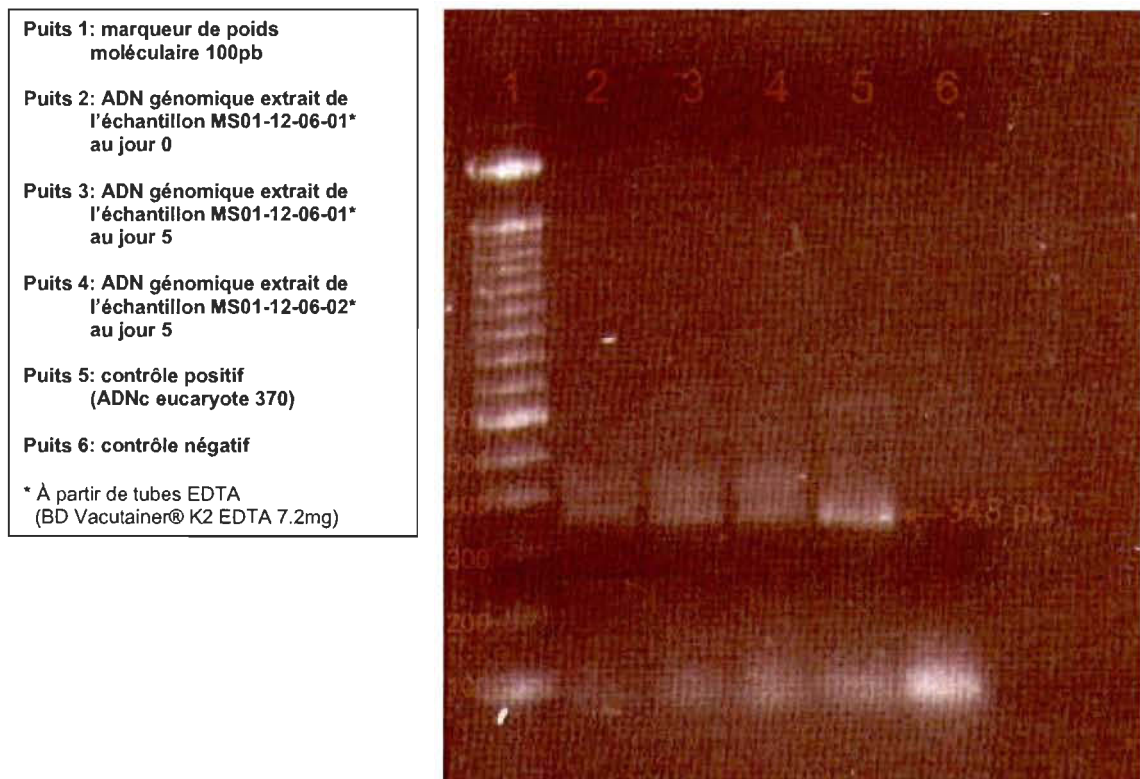


FIGURE 3.1 : Photographie du gel représentant la migration des produits PCR obtenus avec les amorces β -actine sur l'ADN génomique extrait du sang au jour 0 et 5.

La taille attendue est de 348pb. Le puits 5 contient le contrôle positif (ADNc humain) et la taille de l'amplicon est de 348pb. Pour les différentes extractions, les produits PCR ont également une taille de 348 pb. Une seconde bande est observée aux

environs de 420pb pour les puits contenant les 3 extractions (puits 2 à 4). Pour ce qui est du contrôle négatif, aucun produit n'a été amplifié et détecté.

L'extraction d'ADN génomique total des échantillons sanguins s'est avérée efficace puisqu'il y a présence de β -actine dans la totalité des échantillons. Cette protéine est constitutive chez l'homme [46] et puisque l'extraction d'ADN génomique s'effectue à partir du sang humain, il s'agit d'une cible idéale. Les réactifs n'ont pas été contaminés puisque le contrôle négatif ne présente aucune amplification. La présence de bandes supplémentaires à la hauteur de 420pb peut s'expliquer par la présence d'introns dans les échantillons comparativement au contrôle positif qui n'en possède pas, puisqu'il s'agit d'ADN complémentaire [47].

3.2 VÉRIFICATION DE L'EFFICACITÉ DES AMORCES

Après la vérification de l'efficacité de l'extraction, les amorces ont également été testées pour vérifier si elles s'hybridaient bien à de l'ADN génomique bactérien. Puisque ces amorces sont universelles, elles devraient permettre d'amplifier des séquences provenant de diverses bactéries. Pour ce faire, des tests ont été effectués sur de l'ADN génomique bactérien provenant de différentes bactéries connues et inconnues : (a) *E. coli* DH5 α , (b) *E. coli* EC100, *Actinobacillus pleuropneumoniae* 4074 (c) extraction 1 et (d) extraction 2 (*App*4074 #1 et #2) et 2 souches bactériennes inconnues nommées (e) H2 et (f) H3. Les réactifs utilisés sont ceux présentés au tableau 2.5 (réaction 1) et les paramètres sont ceux au tableau 2.6 (paramètres 1). La figure 3.2 présente les résultats obtenus pour la première réaction.

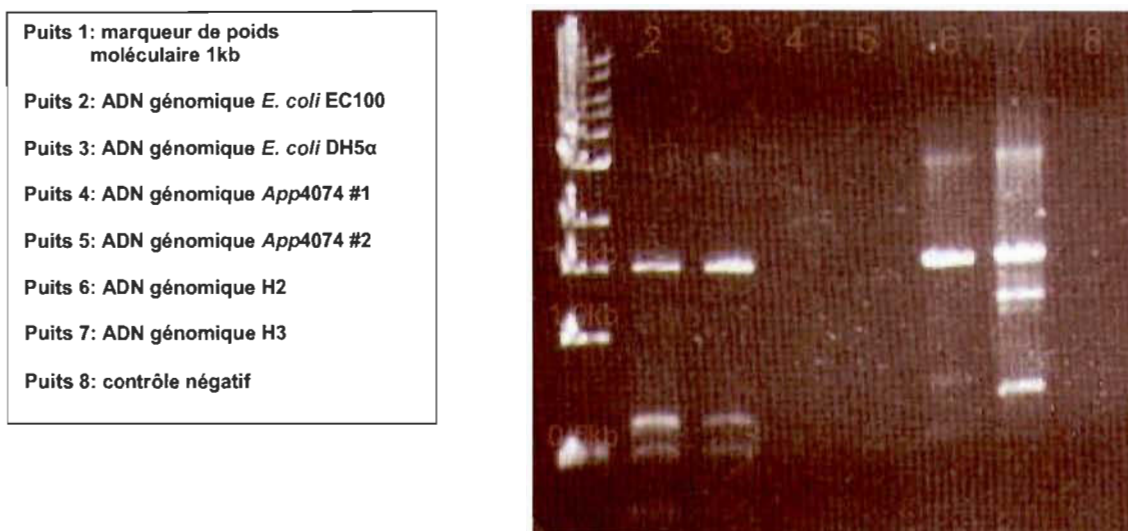


FIGURE 3.2 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique de différentes bactéries avec une température d'appariement des amorces de 50°C.

Le contrôle négatif ne démontre aucune amplification. Il n'y a pas de contrôle positif puisque la totalité des échantillons sont des contrôles positifs. Le gène codant pour l'ARNr16S est présent chez la totalité des procaryotes donc il devrait y avoir une amplification pour chaque souche bactérienne. Par contre, dans ce cas-ci, les deux souches d'*App4074* n'ont pas été amplifiées. Les 4 autres souches possèdent des bandes aux hauteurs de 1.5kb, soit la taille désirée. Il y a également présence de bandes non spécifiques.

La présence de bandes non spécifiques peut s'expliquer par le fait que la température d'appariement théorique est de 60°C, soit 10°C au dessus de la température utilisée dans cette réaction. Les amorces s'apparient à des régions partiellement complémentaires du brin d'ADN. Puisque le but de cette expérience est de détecter la présence d'ADN bactérien, la présence de bandes non spécifiques n'empêche pas la mise en évidence d'ADN bactérien dans les puits 2, 3, 6 et 7. L'absence d'amplification pour *App4074* a été expliquée par une température d'appariement des amorces trop élevée. Une seconde réaction (mêmes réactifs que précédemment) à une température de 48°C (paramètres 2 présentés au tableau 2.6) a été effectuée sur les mêmes souches bactériennes (figure3.3)

Puits 1: marqueur de poids moléculaire 1kb
Puits 2: ADN génomique <i>App4074</i> #1
Puits 3: ADN génomique <i>E. coli</i> DH5 α
Puits 4: ADN génomique H2
Puits 5: ADN génomique H3
Puits 6: ADN génomique <i>E. coli</i> EC100
Puits 7: contrôle négatif



FIGURE 3.3 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique de différentes bactéries avec une température d'appariement des amorces de 48°C.

Le contrôle négatif ne démontre aucune amplification donc il n'y a pas eu de contaminations lors des manipulations et les réactifs n'ont pas été contaminés. *App4074* présente une amplification faible de 1.5kb. Les 3 puits suivants possèdent aussi une bande de 1.5kb en plus de bandes non spécifiques. Le puits 6, contenant la réaction effectuée avec la bactérie *E. coli* EC100, n'a pas d'amplification de 1.5kb. L'ADN génomique était dégradé ou de moins bonne qualité puisqu'il y a eu une amplification avec le même ADN dans la réaction précédente et que deux nouvelles extractions d'ADN génomique ont permis d'obtenir une bonne amplification (figure 3.4) avec les mêmes paramètres que ceux utilisés pour la réaction présentée en figure 3.3. En conclusion, la température d'appariement des amorces qui semble permettre la meilleure réaction est 48°C puisqu'elle a permis d'amplifier le gène codant pour l'ARNr 16S chez la totalité des souches bactériennes testées.

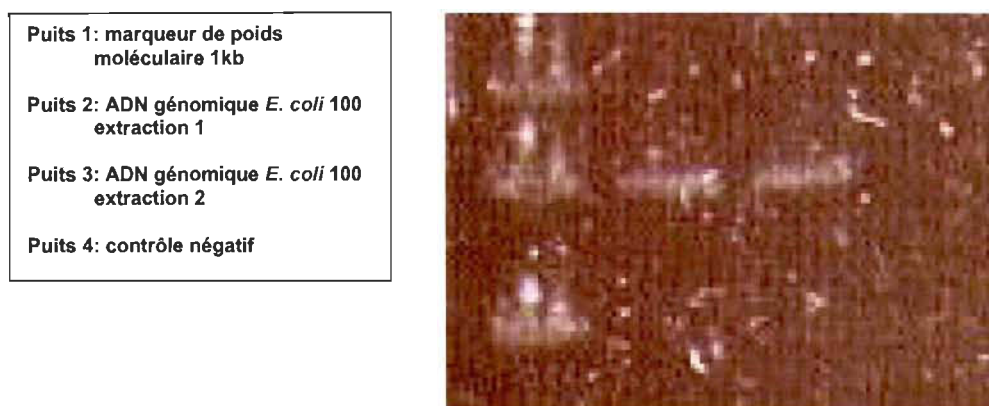


FIGURE 3.4 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique extrait de *E. coli* 100.

3.3 DÉTECTION D'ADN BACTÉRIEN DANS LE SANG PAR RÉACTION PCR

Pour détecter de l'ADN bactérien dans le sang veineux des différents sujets, des réactions PCR ont été effectuées à partir de l'ADN génomique total extrait. Il est important de mentionner que plusieurs réactions ont été effectuées avec chaque échantillon de sang provenant des différents donneurs, mais seulement les résultats obtenus pour un donneur seront présentés. Ceci est dû au fait que les résultats sont les mêmes.

3.3.1 PCR avec différents volumes d'ADN génomique et d'amorces

La première réaction PCR (figure 3.5) a été effectuée avec les paramètres présentés au tableau 2.6 (paramètres 2) et les différents réactifs sont ceux du tableau 2.5 (réaction 1). Le volume d'ADN génomique qui a été utilisé est donc 2µl et 1µl pour les amorces.

Puits 1: marqueur de poids moléculaire 1kb
Puits 2: ADN génomique extrait de l'échantillon MS01-12-06-01* au jour 0
Puits 3: ADN génomique extrait de l'échantillon MS01-12-06-01* au jour 5
Puits 4: ADN génomique extrait de l'échantillon MS01-12-06-02* au jour 5
Puits 5: contrôle positif (ADN génomique <i>E.coli</i> DH5α)
Puits 6: contrôle négatif
* À partir de tubes EDTA (BD Vacutainer®K2 EDTA 7.2mg)*

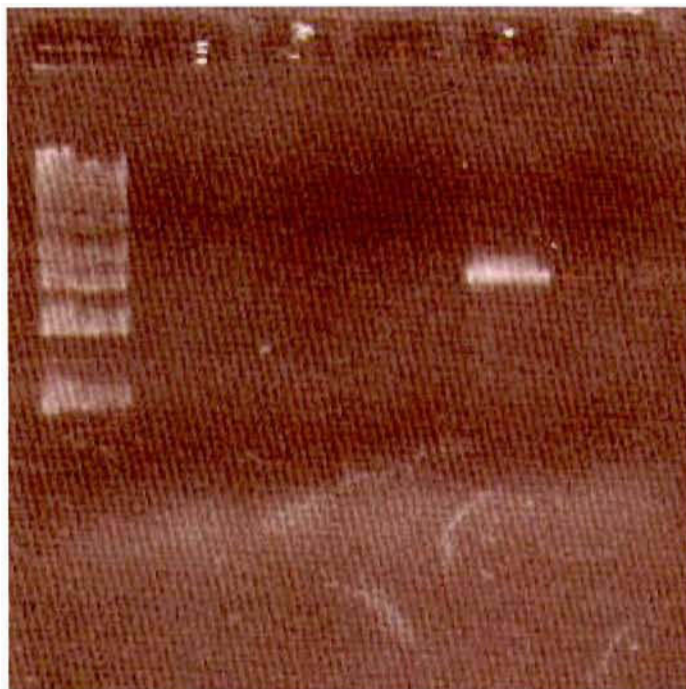


FIGURE 3.5 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique (2µl) avec une température d'appariement des amorces de 48°C.

La taille du produit PCR recherché est de 1505pb. Le contrôle positif, effectuée avec de l'ADN génomique bactérien, a permis d'amplifier un produit PCR de 1.5kb. Pour les 3 extractions, aucune amplification n'a été détectée. Le contrôle négatif n'a, également, aucun produit PCR détecté, donc pas de contaminations des réactifs.

Dans la seconde réaction, la quantité d'ADN génomique a été augmentée pour optimiser la détection d'ADN bactérien. Puisque la détection d'ADN bactérien s'effectue à partir d'ADN majoritairement eucaryote, l'augmentation de la quantité d'ADN au départ augmente la possibilité de détecter de l'ADN bactérien dans les échantillons. Le volume d'ADN génomique utilisé pour la réaction PCR (figure 3.6) a été augmenté à 5µl et celui des amorces à 2.5µl (réaction 2, tableau 2.5). Les paramètres utilisés sont les mêmes que précédemment.

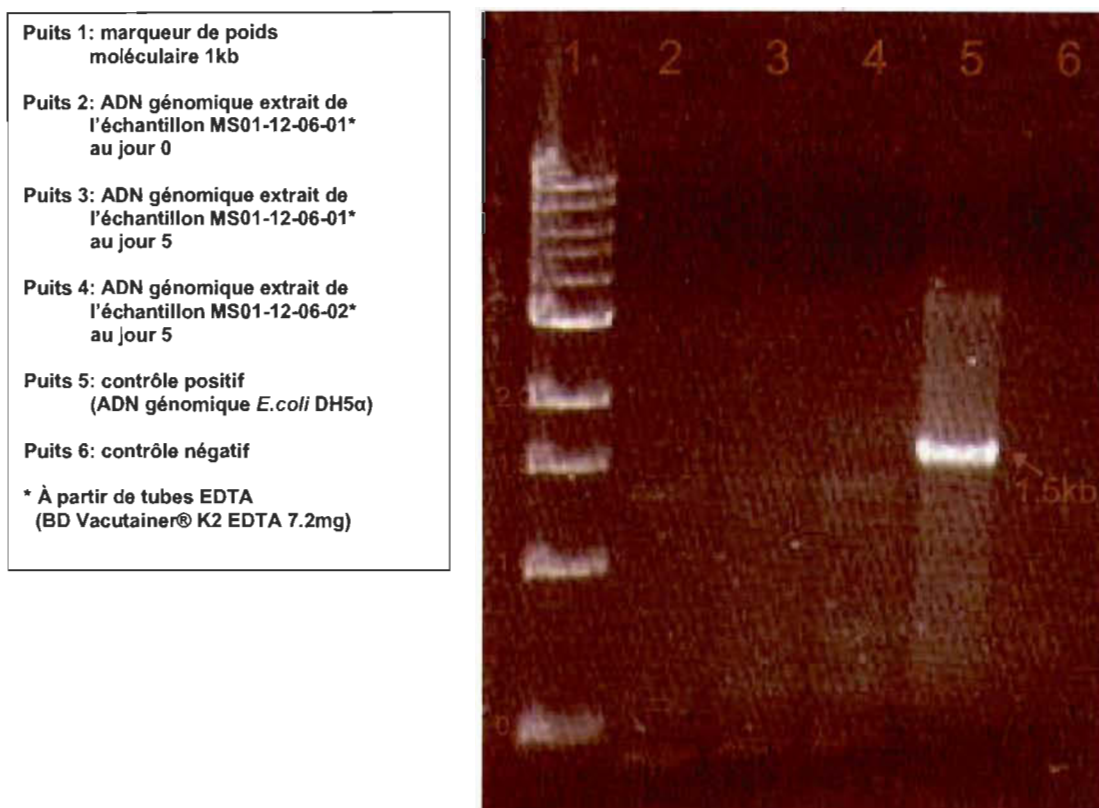


FIGURE 3.6 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique (5µl) avec une température d'appariement des amorces de 48°C.

Le contrôle positif présente une amplification de 1.5kb. Le contrôle négatif n'a aucune amplification. Pour ce qui est des 3 échantillons, ils présentent de faibles amplifications légèrement inférieures à 1.5 kb et 0.5kb.

L'augmentation du volume d'ADN a permis d'amplifier quelques bandes, mais pas celle recherchée d'une longueur de 1.5 kb. Les amplifications observées pour les 3 échantillons sont non spécifiques puisque les amorces amplifient habituellement une séquence de 1.5kb. Les amorces doivent se lier non spécifiquement à des régions partiellement complémentaires sur l'ADN eucaryote [48].

3.3.2 PCR avec différentes températures d'appariement des amorces

En diminuant la température d'appariement des amorces, celles-ci se fixent moins spécifiquement, mais la bande d'intérêt (1.5kb), si elle est amplifiée, peut ensuite être isolée pour le séquençage. Pour ce faire, la prochaine réaction (figure 3.7) a été effectuée à une température de 45°C (paramètres 3 dans le tableau 2.6) et les réactifs sont demeurés les mêmes.

Puits 1: marqueur de poids moléculaire 1kb
Puits 2: ADN génomique extrait de l'échantillon MS01-12-06-01* au jour 0
Puits 3: ADN génomique extrait de l'échantillon MS01-12-06-01* au jour 5
Puits 4: ADN génomique extrait de l'échantillon MS01-12-06-02* au jour 5
Puits 5: contrôle positif (ADN génomique <i>E.coli</i> DH5α)
Puits 6: contrôle négatif
* À partir de tubes EDTA (BD Vacutainer®K2 EDTA 7.2mg)



FIGURE 3.7 : Photographie du gel représentant la migration des produits sur l'ADN génomique (5µl) avec une température d'appariement des amorces de 45°C.

Le contrôle positif présente une amplification de 1.5kb. Le contrôle négatif n'a aucune amplification. Pour ce qui est des trois échantillons, tous possèdent des amplifications légèrement inférieures à 1.5kb et 0.5kb, comme précédemment. Par contre, les échantillons au jour 5 présentent d'autres bandes se situant au-dessus de 0.5kb (≈600pb et ≈700) et une autre à ≈1200pb.

3.3.3 PCR avec une augmentation de la durée de dénaturation

L'ADN génomique est généralement compact et l'étape de dénaturation permet de séparer les brins d'ADN [49]. En augmentant la durée de dénaturation, l'ADN serait plus accessible aux amorces pour permettre un meilleur appariement. Pour cette réaction (figure 3.8), les réactifs sont demeurés les mêmes et les paramètres utilisés sont ceux présentés au tableau 2.6 (paramètre 4). Le temps de dénaturation est passé de 30 à 45 secondes.

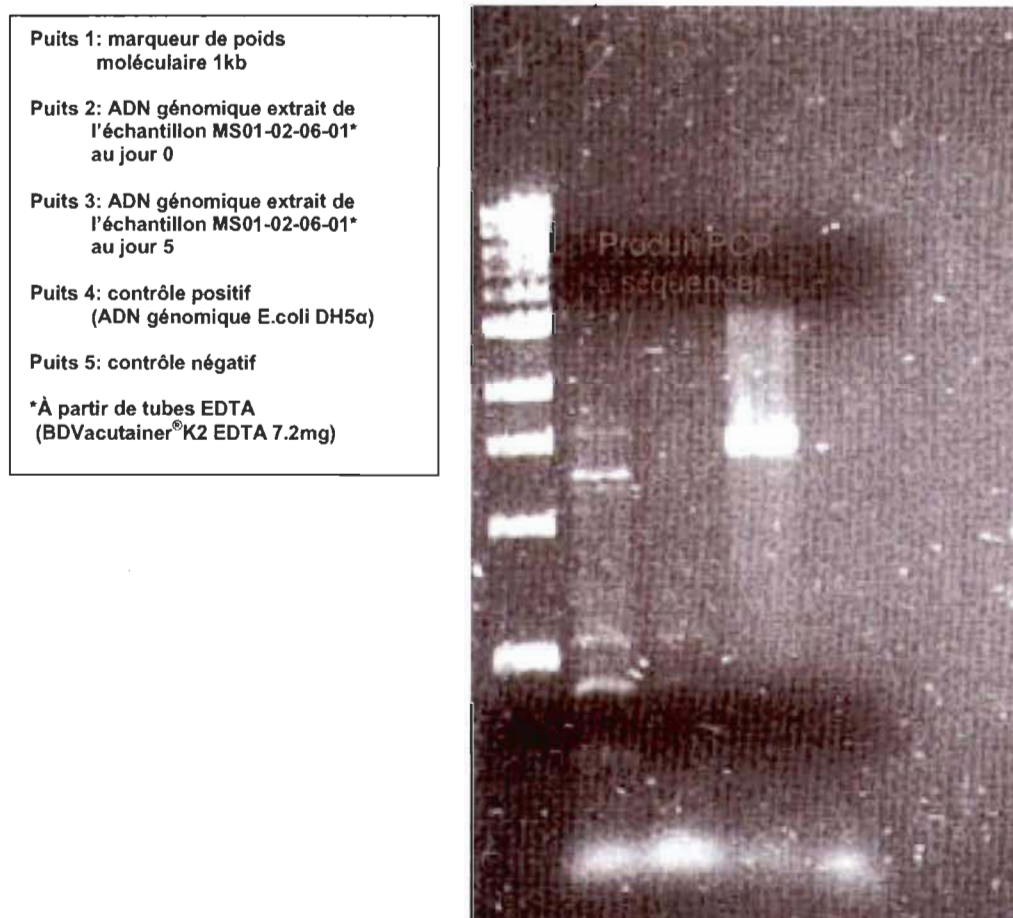


FIGURE 3.8 : Photographie du gel représentant la migration des produits PCR sur l'ADN génomique (5µl) avec une durée de dénaturation de 45s.

Le contrôle positif présente une amplification de 1.5 kb. Le contrôle négatif ne présente aucune amplification. L'échantillon au jour 0 présente 4 bandes distinctes d'une valeur approximative de : 0.4kb, 0.6kb, 1.3kb et 1.55kb. L'échantillon au jour 5 présente

seulement les bandes de $\approx 0.4\text{kb}$ et 0.6kb . La taille du gène codant pour l'ARNr 16S mesure plus ou moins 1540pb selon les espèces [27]. La bande de 1.55kb a donc été clonée pour permettre son séquençage et son identification.

3.4 CLONAGE DU PRODUIT PCR DANS LE VECTEUR pDRIVE

Pour permettre le séquençage, le produit PCR a été cloné dans un vecteur. Puisque l'amplification s'est effectuée à partir d'ADN génomique totale provenant du sang humain et que la possibilité que le produit PCR soit amplifié à partir de différentes bactéries inconnues, le séquençage n'a pas pu se faire à partir du produit PCR uniquement. C'est pourquoi le vecteur pDRIVE a été utilisé pour permettre de sélectionner différents clones et ensuite vérifier leur identité par séquençage.

3.4.1 Électroporation de la ligation dans la bactérie QIAGEN EZ

Les produits PCR purifiés, provenant de l'échantillon MS01-02-06-01 au jour 0, ont été clonés dans le vecteur pDRIVE à l'aide d'une ligation. Celle-ci a été électroporée pour permettre la sélection des bactéries contenant le vecteur recombinant, soit les colonies blanches présentes sur le milieu sélectif LB ampicilline ($100\mu\text{g/ml}$), X-Gal ($80\mu\text{g/ml}$) et IPTG ($50\mu\text{M}$). Les résultats obtenus suite à l'électroporation sont présentés au tableau 3.1.

TABLEAU 3.1
Résultats de l'électroporation de la ligation
des produits PCR et du vecteur pDRIVE

Produit électroporé	Volumeensemencé ¹	Colonies blanches	Colonies bleues
ligation	200 μl	18	3
	[ϵ]	88	29
Contrôle positif (pUC18)	100 μl	-----	TNPC
Contrôle négatif	100 μl	-----	-----

1 : Sur le milieu sélectif LB ampicilline ($100\mu\text{g/ml}$), X-Gal ($80\mu\text{g/ml}$) et IPTG ($50\mu\text{M}$)

[ϵ] : concentration des cellules avant l'ensemencement; TNPC : trop nombreuses pour compter

Le contrôle négatif n'a eu aucune croissance donc il n'y a pas eu de contaminations. Le contrôle positif a été effectué avec un vecteur, pUC18, qui possède le gène codant pour la β -galactodidase, donc les colonies deviennent bleues sur ce milieu sélectif. Les réactifs présents dans le milieu étaient efficaces puisqu'il y a eu une bonne croissance du contrôle positif et que les colonies présentes étaient bleues. Pour la ligation, un premier volume de 200 μ l a étéensemencé et le reste a été concentré pour augmenter le nombre de cellules sur un p \acute tri. La transformation s'est avérée efficace puisqu'il y a des colonies pour les 2 ensemencements. Le rapport de colonies blanches sur le nombre de colonies totales est de 86% pour le premier ensemencement (200 μ l) et de 75% pour le deuxième. La proportion habituelle de recombinants contenant l'insert de PCR après la transformation avec le vecteur de clonage pDRIVE est de 90% pour les cellules compétentes QIAGEN EZ. Les valeurs obtenues sont légèrement inférieures. L'ensemencement de 200 μ l a permis un meilleur rendement (86%) puisqu'il se rapproche plus de la valeur théorique (90%).

Pour sélectionner le produit PCR recherché (1.55kb) en vérifiant la taille de l'insert, un «cracking» a été effectué à partir de patchs effectuées sur 19 colonies blanches sélectionnées aléatoirement. Les résultats obtenus sont présentés à la figure 3.9.

Puits 1: marqueur de poids
moléculaire «supercoiled»
Puits 2 à 20: clones #1 à #19

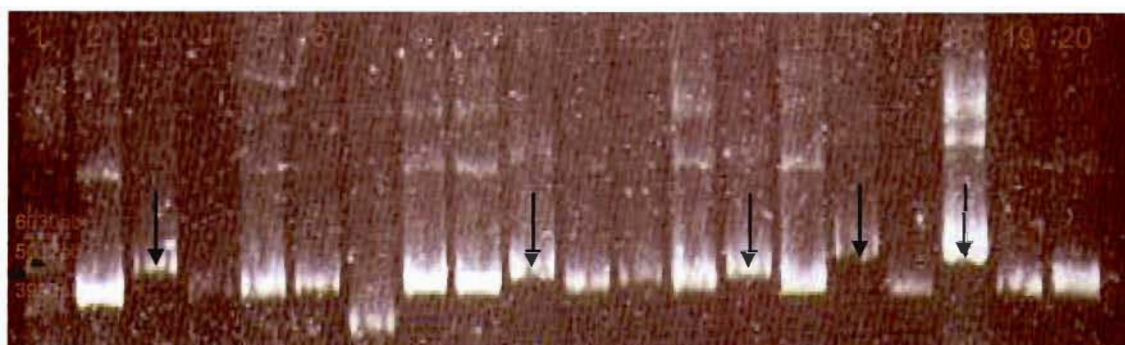


FIGURE 3.9 : Migration du cracking effectué sur 19 clones recombinants.

Puisque la taille du vecteur est de 3.85kb, la taille recherchée du vecteur recombinant est d'approximativement 5.35kb. Seulement deux clones ont cette taille soient les clones 15 et 17 présents dans les puits 16 et 18. Les autres clones correspondent aux autres bandes amplifiées de $\approx 0.4\text{kb}$, 0.6kb et 1.3kb . Les puits 4, 5, 6, 8, 9, 11, 13, 15, 19 et 20 contiennent les vecteurs contenant la bande de $\approx 0.4\text{kb}$. Le puits 12 contient le vecteur contenant la bande de $\approx 0.6\text{kb}$ et les puits 3, 10 et 14 contiennent ceux contenant la bande de $\approx 1.3\text{kb}$. Le puits 7 semble contenir uniquement le vecteur pDRIVE. Par contre, il s'agit peut être de l'insertion de dimères d'amorces puisque l'insertion de 40pb ne serait pas différentiable sur ce gel. L'insertion du dimère a empêché la production de β -galactodidase.

Le clone 17 a été purifié à l'aide d'une mini-prep et il a été envoyé séquençer pour permettre d'identifier l'origine de la séquence insérée. Un *BLAST* dans la banque génomique «Microbes» (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi), qui comprend les bactéries et les archéobactéries, a été effectué sur la séquence obtenue (annexe A) suite au séquençage. Aucune similarité n'a été établie avec aucun génome d'origine procaryote. Un autre *BLAST* a été effectué dans la banque génomique «Human» (www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=9606) pour vérifier si la séquence est d'origine humaine. Une similarité de 98% a été établie avec une séquence se situant sur le chromosome 19 qui contient le gène codant pour une protéine aux doigts de zinc.

3.5 CONCLUSION

Le premier objectif étant de détecter de l'ADN bactérien dans le sang veineux de différents sujets à l'aide de différentes technologies biomoléculaires, on peut conclure que la PCR conventionnelle n'a pas permis d'atteindre le premier objectif puisque aucun produit PCR ne s'est avéré être d'origine bactérienne. La séquence obtenue était d'origine humaine. L'utilisation de différents paramètres n'a pas permis l'amplification de la séquence voulue de 1.5kb. L'amplification d'une séquence humaine peut s'expliquer par la non spécificité de la réaction. La température d'appariement des

amorce était trop basse et a permis l'amplification de plusieurs séquences non spécifiques d'origine humaine. Étant donné que la majorité de l'ADN génomique présent dans les échantillons est d'origine humaine, il n'est pas surprenant d'avoir obtenu ce résultat.

L'absence d'amplification de séquences bactériennes ne signifie pas l'absence d'ADN bactérien dans le sang humain. La PCR conventionnelle n'est pas aussi sensible qu'une PCR quantitative. La première nécessite une quantité d'ADN suffisante pour être observée sur un gel d'agarose ($>500\text{pg}$). Quant à la PCR quantitative, elle peut détecter des concentrations aussi faibles que des femtogrammes [44].

CHAPITRE 4

PCR QUANTITATIVE POUR LA DÉTECTION D'ADN BACTÉRIEN ET MISE EN ÉVIDENCE DE LA VIABILITÉ

La PCR quantitative a été utilisée pour détecter de l'ADN bactérien dans les échantillons sanguins, comme pour la PCR conventionnelle. En plus d'être plus sensible pour la détection, cette technique permet de quantifier le nombre de copies d'ADN. Pour démontrer la viabilité des bactéries dans le sang des différents sujets, la PCR quantitative a été utilisée sur des extractions effectuées au jour 0 et suite à une incubation de 5 jours à 30°C. Il est important de mentionner que dans ce cas-ci, il s'agit d'une PCR semi-quantitative puisque aucune courbe standard n'a été effectuée pour que ce soit quantitatif. Les extractions d'ADN génomique provenaient de sang prélevé dans des tubes contenant aucun anti-coagulant ou différents anti-coagulants, pour vérifier s'il y en avait un qui favorisait ou empêchait la croissance bactérienne.

4.1 PCR QUANTITATIVE SUR DES ÉCHANTILLONS PROVENANT DE TUBES SANS ANTI-COAGULANT

La première réaction a été effectuée avec du sang provenant de deux différents sujets. Les tubes ne contenaient aucun anticoagulant. Les amorces utilisées (tableau 2.8) amplifient également une région du gène codant pour l'ARNr 16S. Dans ce cas-ci, elles amplifient une séquence contenant trois régions hypervariables (V7 à V9). Un contrôle positif a été effectué avec de l'ADN génomique extrait de la bactérie *E. coli* DH5 α (5ng/ μ l). Un contrôle négatif contenant seulement de l'eau Gibco stérile, autoclavée et exposée aux UVs a été utilisé. Les résultats sont exposés à la figure 4.1.

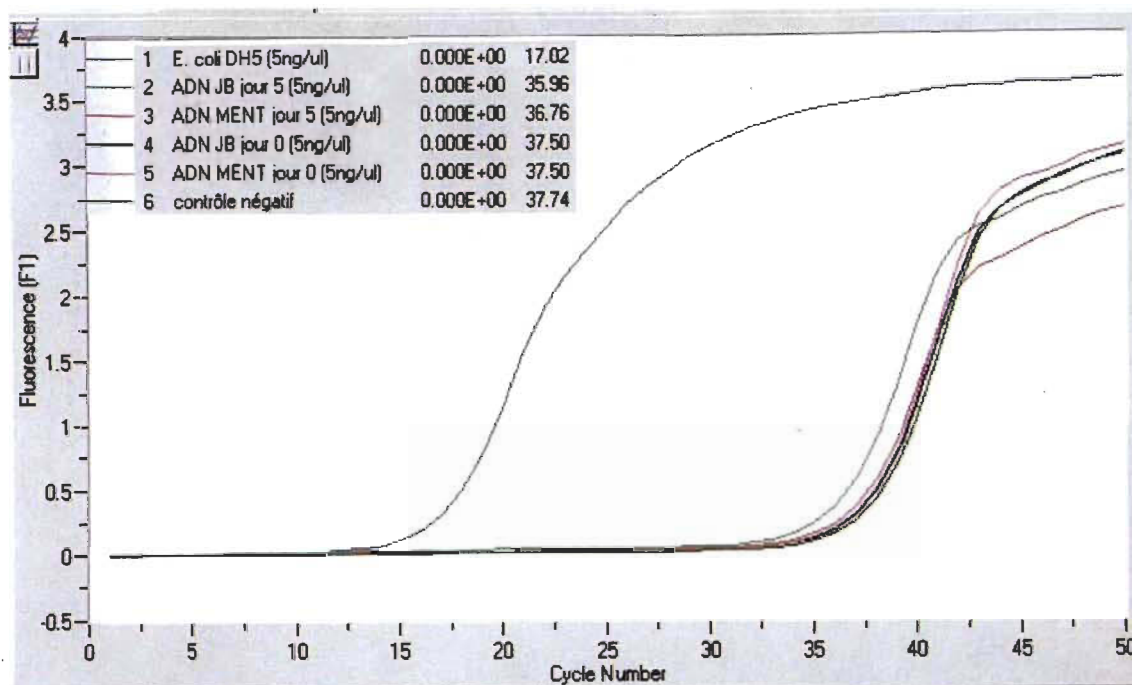


FIGURE 4.1 : PCR quantitative sur 4 extractions provenant de sang contenu dans des tubes sans coagulant de deux sujets au jour 0 et 5.

Le contrôle positif a bien été amplifié comme le démontre la courbe bleue à la figure 4.1. Le contrôle négatif démontre également une amplification qui débute après 37.74 cycles. Pour les deux échantillons au jour 0, l'amplification débute à 37.5 cycles, soit après les échantillons au jour 5 : 35.96 cycles pour l'échantillon ADN JB et 36.76 cycles pour l'échantillon ADN MENT.

L'amplification du contrôle négatif est un phénomène souvent observé lors de l'utilisation du PCR quantitatif avec des amorces ciblant un gène présent chez toutes les bactéries comme dans ce cas ci : gène codant pour l'ARNr 16S [50]. Ceci s'explique par le fait que l'enzyme utilisée pour l'amplification (*Taq* polymérase) est produite à partir de bactéries. Les échantillons contenant l'enzyme contiennent tous de l'ADN bactérien résiduel [51]. Pour les différents échantillons, les résultats sont ceux désirés puisque la quantité d'ADN au jour 5 est supérieure à celle dans les échantillons au jour 0 étant donné que l'amplification s'est effectuée plus tôt pour les échantillons au jour 5.

4.2 PCR QUANTITATIVE SUR DES ÉCHANTILLONS PROVENANT DE TUBES CONTENANT DE L'HÉPARINE ET DU CITRATE (1 SUJET)

Cette réaction (figure 4.2) a été effectuée avec des échantillons de sang prélevés sur un seul sujet et ceux-ci provenaient de tubes contenant deux anti-coagulants différents : citrate-sodium et héparine-sodium. La quantité d'ADN génomique a été augmentée à 10ng/μl comparativement à 5ng/μl lors du PCR quantitatif précédent. L'augmentation de la quantité d'ADN de départ permettra peut-être une meilleure différenciation des échantillons d'avec le contrôle négatif. Le contrôle positif est de l'ADN génomique provenant de la même bactérie que précédemment, mais la concentration est de 5pg/μl qui correspond à ≈ 1000 cellules bactériennes.

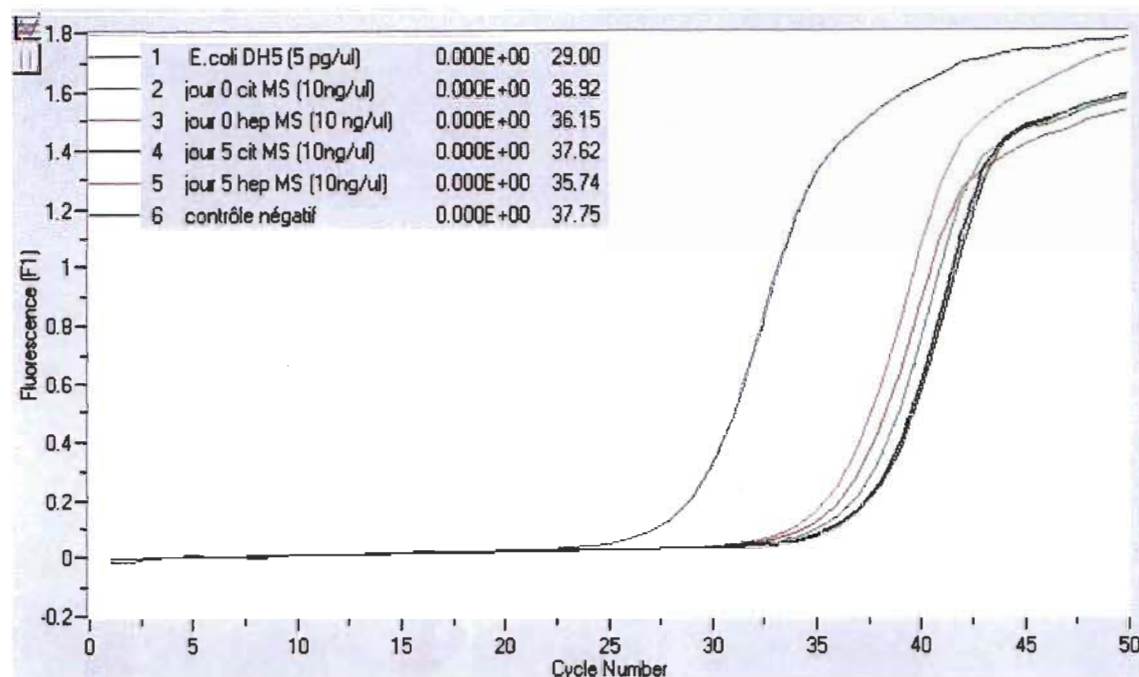


FIGURE 4.2 : PCR quantitative sur 4 extractions provenant de sang contenu dans des tubes avec deux différents coagulants (citrate et héparine) d'un sujet au jour 0 et 5.

Le contrôle positif a bien été amplifié. Le contrôle négatif a été amplifié après 37.75 cycles. Les 4 échantillons ont été amplifiés avant le contrôle négatif. Par contre, l'échantillon cit MS au jour 5 a été amplifié seulement après 37.62 cycles. L'échantillon cit MS au jour 0 a été amplifié plus tôt (36.92). Les échantillons hep MS au jour 0 et jour

5 respectent l'hypothèse selon laquelle une incubation de 5 jours permettrait la croissance bactérienne, puisque l'échantillon au jour 5 contient plus de copies d'ADN que l'échantillon au jour 0. Si l'on compare les 2 anti-coagulants, l'héparine semble être le meilleur pour permettre la croissance bactérienne.

Dans la seconde PCR (figure 4.3), les mêmes échantillons ont été utilisés pour vérifier la reproductibilité des résultats.

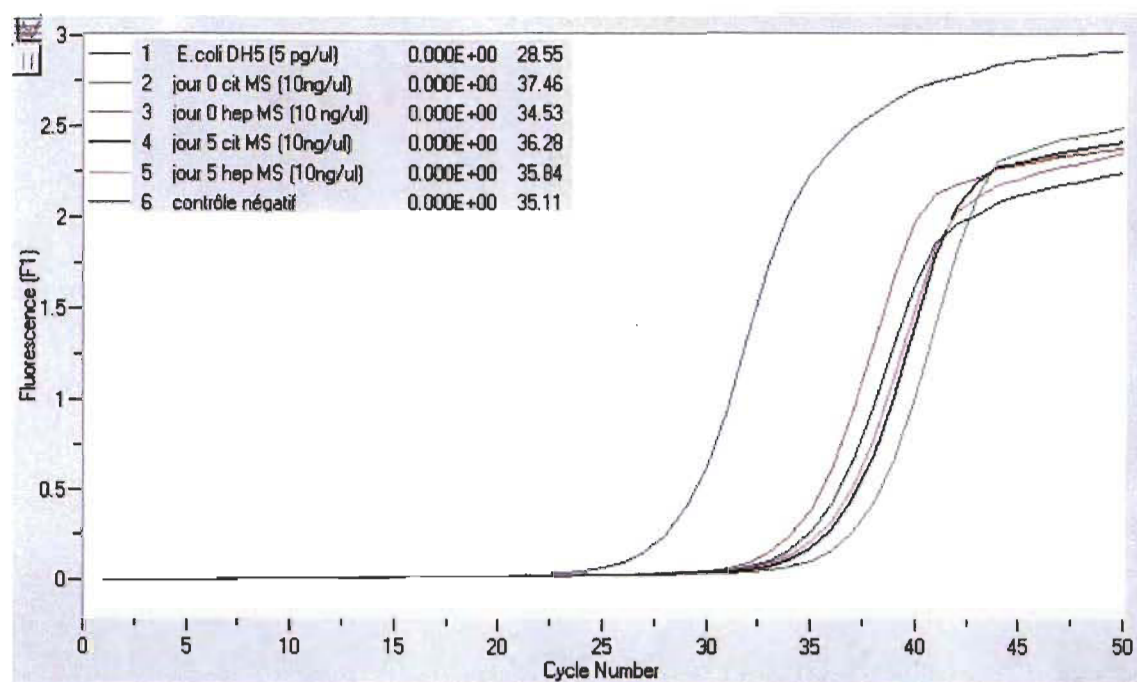


FIGURE 4.3: PCR quantitative sur 4 extractions provenant de sang contenu dans des tubes avec deux différents coagulants (citrate et héparine) d'un sujet au jour 0 et 5 (essai #2).

Le contrôle positif a bien été amplifié. Le contrôle négatif est encore amplifié, mais cette fois-ci, après seulement 35.11 cycles, donc plus tôt que les échantillons (a) cit MS jour 0, (b) cit MS jour 5 et (c) hep MS jour 5.

Avec les mêmes échantillons de même concentration, il n'a pas été possible de reproduire des résultats semblables à ceux obtenus précédemment. Les échantillons n'ont pas de différences significatives comparativement au contrôle négatif puisque

celui-ci est parfois amplifié avant les échantillons, donc la présence d'ADN bactérien serait due à l'ADN résiduel provenant de l'enzyme.

4.3 PCR QUANTITATIVE SUR DES ÉCHANTILLONS PROVENANT DE TUBES CONTENANT DU CITRATE (4 SUJETS)

Une réaction PCR a été effectuée (figure 4.4) avec des échantillons provenant de tubes contenant seulement du citrate de sodium et ce, sur 4 sujets au jour 0 et 5.

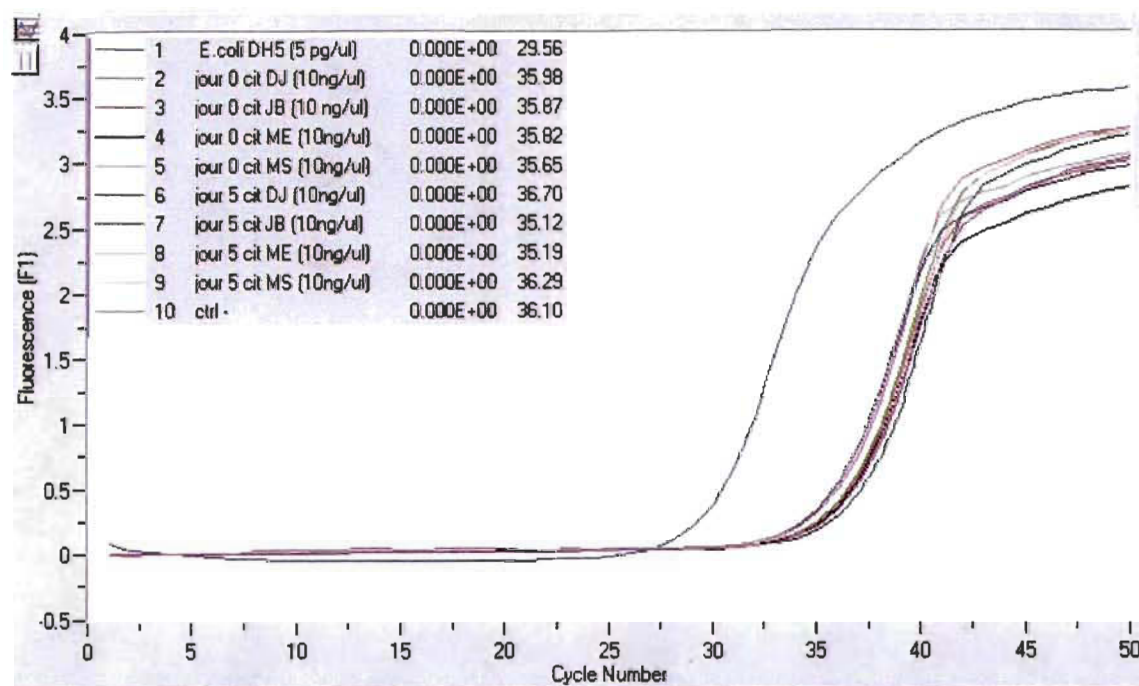


FIGURE 4.4: PCR quantitative sur 8 extractions provenant de sang contenu dans des tubes contenant du citrate de quatre sujets au jour 0 et 5.

Pour cette réaction, le contrôle positif a bien été amplifié. Le contrôle négatif a été amplifié à 36.10 cycles. Deux échantillons contiennent moins d'ADN bactérien que le contrôle négatif : cit DJ et cit MS au jour 5. Les échantillons cit JB et cit ME au jour 5 sont amplifiés avant les échantillons au jour 0. Il y aurait donc plus d'ADN bactérien dans les échantillons au jour 5 qu'au jour 0.

4.4 PCR QUANTITATIVE SUR DES ÉCHANTILLONS PROVENANT DE TUBES CONTENANT DE L'HÉPARINE (3 SUJETS)

Un autre anti-coagulant testé est l'héparine-sodium et cette fois-ci, sur 3 sujets différents. La figure 4.5 présente les résultats obtenus pour la réaction effectuée avec cet anti-coagulant.

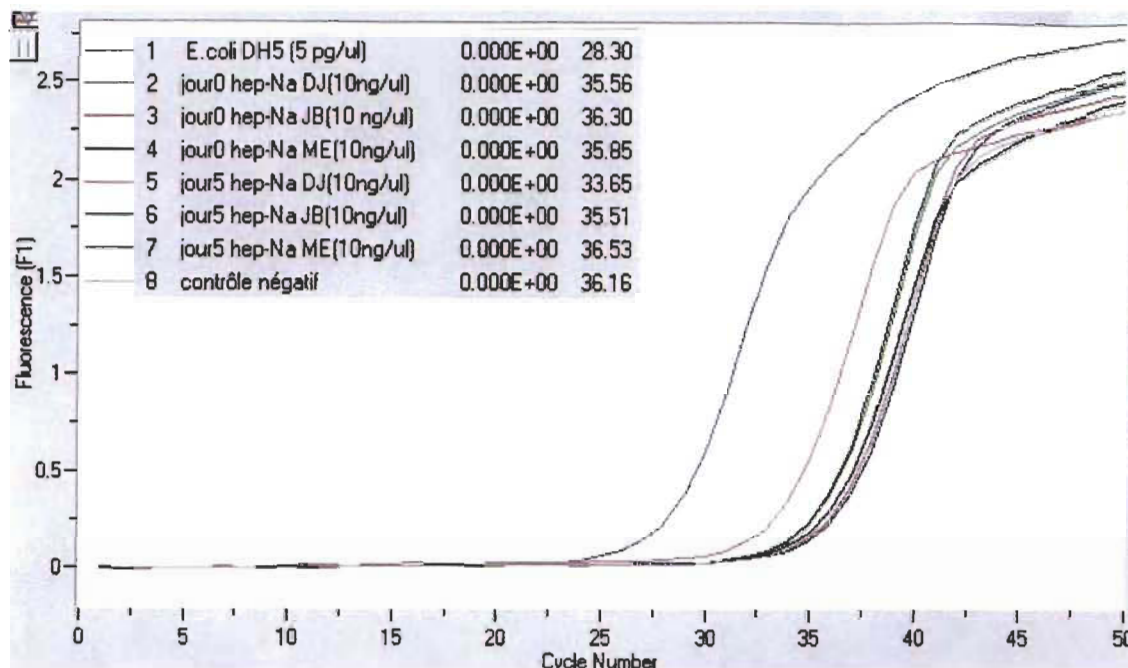


FIGURE 4.5: PCR quantitative sur 3 extractions provenant de sang contenu dans des tubes de l'héparine-sodium de trois sujets au jour 0 et 5.

Le contrôle positif a bien été amplifié. Le contrôle négatif a été amplifié après 36.16 cycles. Deux échantillons ont été amplifiés après le contrôle négatif : hep-Na JB au jour 0 et hep-Na ME au jour 5. L'échantillon hep-Na DJ respecte l'hypothèse selon laquelle il y aurait une croissance bactérienne suite à une incubation de 5 jours puisque l'échantillon au jour 5 est amplifié avant l'échantillon au jour 0 et ce dernier, est amplifié avant le contrôle négatif.

4.5 CONCLUSION

L'utilisation de la PCR quantitative avait pour but de détecter de l'ADN bactérien dans le sang veineux de différents sujets sains (premier objectif) et il permettait également de déterminer préliminairement (les résultats n'étant pas quantitatifs, mais seulement semi-quantitatifs) la croissance bactérienne suite à une incubation de 5 jours, répondant ainsi au second objectif qui était de démontrer la viabilité des bactéries. Certains échantillons ont démontrés une tendance présumant la présence d'ADN bactérien dans le sang et même la croissance bactérienne. Par contre, la non reproductibilité des résultats et le facteur d'erreur dû à la présence d'ADN bactérien résiduel dans les échantillons font en sorte que les résultats ne sont pas significatifs.

Un autre point important est que la concentration des échantillons est calculée à partir de l'ADN génomique total extrait des prélèvements sanguins. Un échantillon contenant 10ng/μl ne contient pas nécessairement la même quantité d'ADN bactérien que dans un autre échantillon de même concentration. Pour éliminer ce facteur, une courbe standard effectuée avec des amorces ciblant un gène constitutif chez l'humain, par exemple la β -actine, permettrait d'éliminer le facteur d'erreur dû à la présence d'ADN eucaryote et ainsi de normaliser tous les échantillons.

CHAPITRE 5

ARTICLE SCIENTIFIQUE

5.1 Interaction of human sequences with the bacterial RNA polymerase in *Escherichia coli* revealed by plasmid-vector promoter trapping

Marie-Edith Nepveu-Traversy¹, Maurice Boissinot², Éric Asselin¹ and Marc Sirois^{1*}

¹Laboratoire de génomique bactérienne
Département de Chimie-Biologie
Université du Québec à Trois-Rivières
3351, boulevard des Forges
Trois-Rivières, Québec
Canada G9A 5H7

²Molecular Diagnostics
Centre de Recherche en Infectiologie de l'Université Laval
Centre hospitalier universitaire de Québec, Pavillon CHUL
2705 Boulevard Laurier, RC-709
Sainte-Foy, Québec
Canada G1V 4G2

* Corresponding author

Tel. : 819-376-5011 ext. 3999

Fax: 819-376-5084

Email: marc.sirois@uqtr.ca

5.2 RÉSUMÉ

Lors de recherches, ayant pour but la détection d'ADN bactérien dans le sang humain, une technique, qui permet l'insertion de séquences promotrices bactériennes, a révélé l'interaction de l'ARN polymérase d'*E. coli* avec des séquences humaines. Fait étonnant, étant donné que le mécanisme de transcription entre les procaryotes et les eucaryotes diffère considérablement. De plus, ces séquences possèdent une similarité avec les motifs consensus présents dans les promoteurs bactériens. Ceci pourrait expliquer partiellement l'interaction avec la polymérase bactérienne. Les séquences humaines ayant une activité promotrice se retrouvent dans des régions non codantes du génome humain. Ces régions dites non codantes, tels les introns, possèdent possiblement une fonction qui est encore inconnue jusqu'à ce jour. La complexité du génome humain suppose qu'il y a effectivement des interactions que l'on ne soupçonne guère.

5.3 ABSTRACT

Transcription mechanism has been extensively studied since several years. Many aspects are well understood but some remains obscures. Transcription between different organisms is known to possess some similarities, but several differences too. Eukaryote promoting activity with bacterial machinery has never been considered before. Surprisingly, eukaryote sequences seem to have a promoting activity in bacterial cells, if we base our fundament on this study. It was unexpected considering that transcription in eukaryote cells is more complex than in bacterial cells. These human sequences possess regions that are similar to bacterial consensus motifs. All human sequences correspond to non-coding DNA in human genome. This finding doesn't tell us why human sequences have promoting activity in bacterial cells, but it reveals us the complexity of the genome.

5.4 INTRODUCTION

While investigating on the presence of bacterial DNA in the blood of healthy subjects by using a promoter trapping technique, we found that eukaryotic sequences, more precisely human sequences, possess promoting activity with the *E. coli* RNA Polymerase (RNAP). At first, this technique was supposed to reveal the presence of bacterial DNA with promoting activity. This result was unexpected considering the differences between bacterial and human promoters.

The transcription process begins with the binding of RNAP to DNA in the promoter region near the Transcription Start Site (TSS). How RNAP recognizes this specific region among this great amount of DNA is not well understood to date and remains a field of investigation [1]. It is known that promoter regions possess some sequence motifs located specifically in the vicinity of the TSS. For example, the bacterial promoter contains at least three RNAP recognition sequences: the -10 element, the -35 element and the UP element [2, 3]. The eukaryote promoter contains consensus sequences as well, but it is generally more complex than the bacterial promoter. Usually, the eukaryote promoter contains three common consensus sequences named TATA box, the initiator element (Inr) and the Downstream Promoter Element (DPE) that are located between -35 and +35 relative to the TSS [4]. Furthermore, there are many other consensus motifs that regulate transcription like TFIIB Recognition Element (BRE), CCAAT-box and GC-box [5].

Consensus motifs are important in the regulation of transcription but other factors are also involved in this process: DNA stability, curvature and bendability are known to be different in the region surrounding the promoter than in the coding region. First, DNA stability is known to be lower upstream the TSS than downstream due to the higher AT content and this for both prokaryote and eukaryote promoters [1]. Secondly, DNA curvature upstream the promoter sequence has been observed for both promoters [6, 7], but the magnitude of the curvature for the eukaryote promoter seems to be less pronounced compared to the bacterial one. Both kinds of promoters reveal common

pattern for bendability, which also plays an important role in transcription. In fact, the similarities between eukaryote and prokaryote promoters consist mainly in the stability and bendability of DNA. However, their nucleotide composition and curvature patterns seem to be very different [1]. Despite the fact that both types of promoters share some similarities, the finding of eukaryote DNA promoting activity in *Escherichia coli* cells remained surprising.

Interaction between sequences from a particular organism and the RNAP of another organism has been seen before. For example, a study demonstrated that DNA from *Streptococcus pneumoniae* and other organisms rich in AT nucleotides had an effect on promoting activity when cloned in a vector expressed in *Escherichia coli* [8]. Nevertheless, the interaction between human DNA and bacterial RNAP has never been studied before. Our data suggest that certain human sequences can interact with the bacterial machinery that direct mRNA synthesis.

5.5 METHODOLOGY

5.5.1 Plasmid construction

Plasmid-vector promoter trapping allows the insertion of DNA in a specific plasmid that possesses a reporter gene (*gfp*). Instead of containing a promoter upstream, it contains a polycloning site. The insertion of a promoting DNA sequence causes the expression of the protein, in this case, Green Fluorescent Protein (GFP). The plasmid that has been used, pSP72-GFP (Figure 1), is a combination of two plasmids: pSP72 and pGreenTIR. The first one confers the ampicilline resistance and the polycloning site. The GFP cassette was isolated from pGreenTIR (Figure 1) because it contains the enhancer (ENH), the Shine-Dalgarno sequence (SD) and the start codon [9]. To realize the plasmid construction and isolate the GFP cassette, an enzymatic digestion was done with EcoRI (10U/reaction, NEB) on the plasmid pGreenTIR (100ng). The pSP72 plasmid (100ng, Promega) was digested with the same enzyme to permit the incorporation of the GFP

cassette downstream the polycloning site. The pSP72 plasmid and GFP cassette were ligated with T4 ligase (400U/reaction, NEB). As a result, the pSP72-GFP contains a GFP cassette without a promoter upstream and can be selected by the presence of ampicilline resistance.

5.5.2 Enzymatic digestions

Insertion of DNA fragments in pSP72-GFP was possible by using two compatible enzymes: BamHI and Sau3A 1. The plasmid (1500ng) was digested with BamHI (10U/reaction, NEB), which the restriction site is found 49 nucleotides upstream the GFP cassette. The total genomic DNA was isolated from four blood samples with a DNA extraction kit (QIAamp[®] DNA Blood Mini Kit, QIAGEN). Sau3A1 (10U/reaction, NEB) was used to digest the genomic DNA (≈ 350 ng/reaction). Afterwards, T4 ligase (400U/reaction, NEB) was used to produce recombinant plasmids: DNA fragments, generated by enzymatic digestion, were inserted upstream the GFP cassette, in the polycloning site (BamHI restriction site).

5.5.3 Transformation and Selection

To certify the presence of promoting activity among the recombinant plasmids, the ligations (2.5 μ l) were incorporated, with the *E. coli* pulser (Biorad) into electrocompetent cells (40 μ l): *Escherichia coli* DH5 α . Negative control was done with cells only to check ampicilline efficacy in culture media and to certify the absence of contaminants. Another control was done with the plasmid pUC18, with an ampicilline resistance, to make sure reagents were efficient. TSB (1ml) was added to cells after transformation to bring them nutrients to survive. Cells were incubated 30 minutes (37°C) before plating 100 μ l on selective media: Luria-Bertani (LB) with ampicilline (100ng/ μ l). To allow bacterial growth, overnight incubation at 37°C was performed. UVs lamp allowed selection of GFP expression. Fluorescents colonies were selected for following cracking.

To select different and same sizes of inserts, cracking was performed on GFP positives clones. A cracking solution was prepared with 9 volumes of cracking buffer (EDTA 0.5M pH8, sodium dodecyl sulphate (SDS) 0.04M, bromophenol blue 8×10^{-4} M, glycerol 10%) for 1 volume of sodium hydroxide (NaOH 1N). Partial GFP positives colonies were taken with tooth-picks and resuspended in 25µl of sterile water. Next, 25µl of cracking solution were added to the suspension and then filled into the agarosis gel. Migration was performed at 120V for 1 hour. To approximate the size of the plasmids, supercoiled DNA ladder was used. Gel photography was done with Molecular Imager Gel DOC system (BIORAD) after SYBR[®]Safe (10 000X concentrated in diméthyl sulfoside (DMSO), Invitrogen) dying. GFP positives clones were selected for sequencing in function of there sizes: different sizes to check the variability of the inserts and same sizes to verify redundancy.

5.5.4 Sequencing and bioinformatics analysis

Sequencing was performed on GFP positives clones, using the reverse primer pSP72-GFP (5'-GGGACAACTCCAGTGAAAAG- 3'), after DNA purification (QIAprep[®] Miniprep, QIAGEN[®]). Samples were sent to bimolecular analysis plate-form of Laval University (www.bioinfo.ulaval.ca). Results were analysed using *BLAST* (Basic Local Alignment Search Tool), used to compare sequences with genomic banks (human and bacterial) from NCBI (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>).

Sequences were then analysed with another program: *Promsearch* (www.molquest.com). This program predicts the number of promoters in a sequence by determining the presence of consensus motifs. In this case, *bProm* was used to find bacterial promoters in the promoting sequences isolated from recombinant plasmids. The program predicts potential transcription sites position of bacterial genes regulated by sigma70 promoters.

5.6 RESULTS

5.6.1 Transformation efficacy

After electroporation of ligation and plating on selective media, UV selection was done and GFP positive clones were counted to determine the percentage of GFP positives on total clones (table 5.1). The average of GFP positive clones is 8.71 %. Negative controls were all negative on LB amp100 media and all media controls with pUC18 showed confluent growth.

5.6.2 Sequencing results

A total of 69 GFP positives clones were selected among 1538 for sequencing (data shown in annexe B). Given that the first purpose of this experiment was to detect bacterial DNA in human blood, table 5.2 presents the cloned sequences which possess a similarity with bacterial DNA.

Of the 69 clones, 7 had similarity with bacterial DNA sequences. Six species has been identified after sequencing: *Bacillus cereus*, *Eubacterium siraeum*, *Bacteroides capillosus*, *Cyanothece* sp., *Escherichia coli* and *Salmonella enterica*. *Bacillus cereus* is represented 3 times and *Cyanothece* sp. 2 times. The similarity percentage is situated between 79% and 100%. Nucleotidic position represents the position of the identified sequence on the total inserted sequence. For example, clone ME4-11-07 possesses an insert of 418 nucleotides; 158 nucleotides on 171 are similar to the organism *Bacillus cereus* and this identified sequence is found at the position 249-418 of the inserted sequence. Another 50 nucleotides is similar to *Eubacterium siraeum* and is found at position 189-218. A part of the left segment is similar to human DNA and is found at position 14-100 (table 5.3). After sequencing, the majority of the cloned sequences with promoting activity turned out to be similar to human sequences and they are presented in table 5.3. No redundancy was observed among all sequences.

A total of 65 cloned sequences possess a similarity with human sequences. The similarity percentage is located between 91% and 100%. Some cloned sequences represent more than one human sequence. For example, the promoting sequence inserted into clone DJ19-12-07 have a similarity with four different sequences based on the fact that four *Sau3A* digested fragments were ligated together and inserted in the plasmid pSP72-GFP.

All cloned sequences possessed a similarity with non coding sequences. Sequences were located within a gene, but in introns, or it was located between two genes on a specific chromosome. For example, sequence 7B-09-07, who possesses a similarity with the gene that codes for the biliverdin reductase A, corresponds to a non-coding region of the gene. In other cases cloned sequences match human DNA sequences that are located between two genes and are not known to be coding DNA.

5.6.3 Prediction of promoters in cloned sequences

The objective of this experiment was to find bacterial promoting sequences in human blood. Some sequences had a similarity with bacteria but the majority turned out to be human. Bacterial promoters were predicted for each sequence to find out if it possesses the consensus motifs -10 and -35. Because the plasmid is electroporated in bacteria, bacterial RNAP interact usually with bacterial promoters. Table 5.4 presents the predicted bacterial promoters for all cloned sequences.

The majority (62) of the cloned sequences possesses one or two predicted promoters. Only 7 clones have no predicted promoter in their cloned sequences: 6J-10-07, ME3-12-07, ME7-12-07, MS24-12-07, DJ22-12-07, JB11-12-07 and JB13-12-07. Average position of predicted promoters is 190 bases upstream the GFP cassette (annexe B).

Promoters are AT rich sequences [1]. The AT percentage of all sequences was calculated and the average among all sequences was 59.68 % and 40.32 % for GC content.

5.7 DISCUSSION

Promoter trapping technique was used to detect the presence of bacterial DNA in blood samples from healthy subjects. This technique allowed the detection of bacterial DNA because 7 sequences among 69 had similarity with bacterial DNA. These sequences had bacterial consensus motifs that are located in the bacterial promoter (table 5.4). These results were expected taking into account that cloned sequences, with promoting activity, were supposed to be identified as bacterial sequences containing the consensus motifs that can interact with the bacterial RNAP [3]. All other sequences were similar to human DNA. Extraction was done with human blood, which contains eukaryote DNA. The presence of human sequences in plasmid-vector is understandable, but interaction between these sequences and *E. coli* RNAP is beyond belief.

The *BLAST* search showed that 65 sequences had similarity with human DNA. It also reveals that cloned sequences possessed regions similar to bacterial consensus motifs found into bacterial promoter. Only 7 sequences did not have bacterial promoter predicted. Promoting activity would not be only regulated by the presence of consensus motifs. Nucleotidic sequence is important for the polymerase recognition but is not the only factor implicated in transcription. First, the exact consensus motifs are found in only a few promoters [10]. Also, the sequences comprise only 6 to 10 bases that are degenerated [1]. The average promoter conserves only 3.9 bases for the -35 consensus sequence (TTGACA) and 4.2 bases for the -10 consensus sequence (TATAAT). The probability of finding sequences similar to bacterial promoter in regions other than promoter is relatively high. In theory, promoter sequences would occur every 200 bases in a random sequence [11]. For the 7 sequences with no motifs, *E. coli* DH5 α RNAP interacted with sequences even if it wasn't consensus motifs. It is known that polymerase

can have lower affinity for sequences and protein is translated as well. Lower affinity results generally in less protein produced [12]. Fluorescence quantification can answer to this by comparing the fluorescence emitted by bacterial containing plasmid with promoting sequences inserted. Sequences with no consensus motif would emit less fluorescence because the promoter would interact poorly with the sequence.

Transcription is not regulated only by the presence of consensus motifs: DNA conformation has an important role [13]. In this case, plasmid were utilised to realize this experiment and plasmid are circular, coiled or supercoiled [14]. But coiled and supercoiled conformation can correspond to DNA curvature observed in promoter region. If we consider this fact, human introns are more curved than promoter and curvature in *E. coli* promoter is high too [6]. Cloned sequences were found in non coding region like introns. DNA conformation and the presence of consensus motifs in the human sequences inserted in the plasmid can explain transcription of the *gfp* gene. Furthermore, promoters are known to possess a high AT content compared with coding region [1]. Average of AT content for all sequences was 59.68% which is quite high. In contrast to the *E. coli* genomic AT content (49.9%), we observed richer AT sequences for the cloned sequences. When we compared with human genomic content (60.02%), a lower percentage is observed for cloned sequences [15]. In conclusion, AT content is high so it responds to the theory that a promoter is rich in AT bases. Average position of predicted promoters is 190 bases upstream the GFP cassette. Promoter can be situated up to 200 bases upstream the TSS [13].

Data suggests an interaction between human DNA and bacterial transcription mechanism. This has not been reported before. These non coding sequences have probably a role to play that we don't understand yet. Further researches with foot printing could show us the real interaction between *E. coli* RNAP and human sequences. Similarity with bacterial consensus motifs is perhaps due to random, considering that genome is wide, but it can be attributed to a real purpose.

5.8 REFERENCES

1. Kanhere, A. and M. Bansal, *Structural properties of promoters: similarities and differences between prokaryotes and eukaryotes*. Nucleic Acids Res, 2005. 33(10): p. 3165-75.
2. Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. and Gourse, R.L., *A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase*. Science, 1993. 262(5138): p. 1407-13.
3. Estrem, S.T., Ross, W., Gaal, T., Chen, Z.W., Niu, W., Ebright, R.H. and Gourse, R.L., *Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase alpha subunit*. Genes Dev, 1999. 13(16): p. 2134-47.
4. Schumacher, M.A., A.O. Lau, and P.J. Johnson, *Structural basis of core promoter recognition in a primitive eukaryote*. Cell, 2003. 115(4): p. 413-24.
5. Bucher, P., *Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences*. J Mol Biol, 1990. 212(4): p. 563-78.
6. Gabrielian, A.E., D. Landsman, and A. Bolshoy, *Curved DNA in promoter sequences*. In Silico Biol, 1999. 1(4): p. 183-96.
7. Jauregui, R., Abreu-Goodger, C., Moreno-Hagelsieb, G., Collado-Vides, J. and Merino, E., *Conservation of DNA curvature signals in regulatory regions of prokaryotic genes*. Nucleic Acids Res, 2003. 31(23): p. 6770-7.
8. Dillard, J.P. and J. Yother, *Analysis of Streptococcus pneumoniae sequences cloned into Escherichia coli: effect of promoter strength and transcription terminators*. J Bacteriol, 1991. 173(16): p. 5105-9.
9. Miller, W.G. and S.E. Lindow, *An improved GFP cloning cassette designed for prokaryotic transcriptional fusions*. Gene, 1997. 191(2): p. 149-53.
10. Rangannan, V. and M. Bansal, *Identification and annotation of promoter regions in microbial genome sequences on the basis of DNA stability*. J Biosci, 2007. 32(5): p. 851-62.
11. O'Neill, M.C., *Escherichia coli promoters. I. Consensus as it relates to spacing class, specificity, repeat substructure, and three-dimensional organization*. J Biol Chem, 1989. 264(10): p. 5522-30.
12. Brutinel, E.D., Vakulskas, C.A., Brady, K.M. and Yahr, T.R., *Characterization of ExsA and of ExsA-dependent promoters required for expression of the*

Pseudomonas aeruginosa type III secretion system. Mol Microbiol, 2008. 68(3): p. 657-71.

13. Plaskon, R.R. and R.M. Wartell, *Sequence distributions associated with DNA curvature are found upstream of strong E. coli promoters*. Nucleic Acids Res, 1987. 15(2): p. 785-96.
14. Sambrook, J. and D.W. Russell, *Molecular cloning : a laboratory manual*. 3rd ed. 2001, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. 3 v. (various pagings).
15. Horton, H.R., *Principes de biochimie*. 1994, Bruxelles: De Boeck Université. xi, 720 p.

5.9 FIGURES LEGENDS

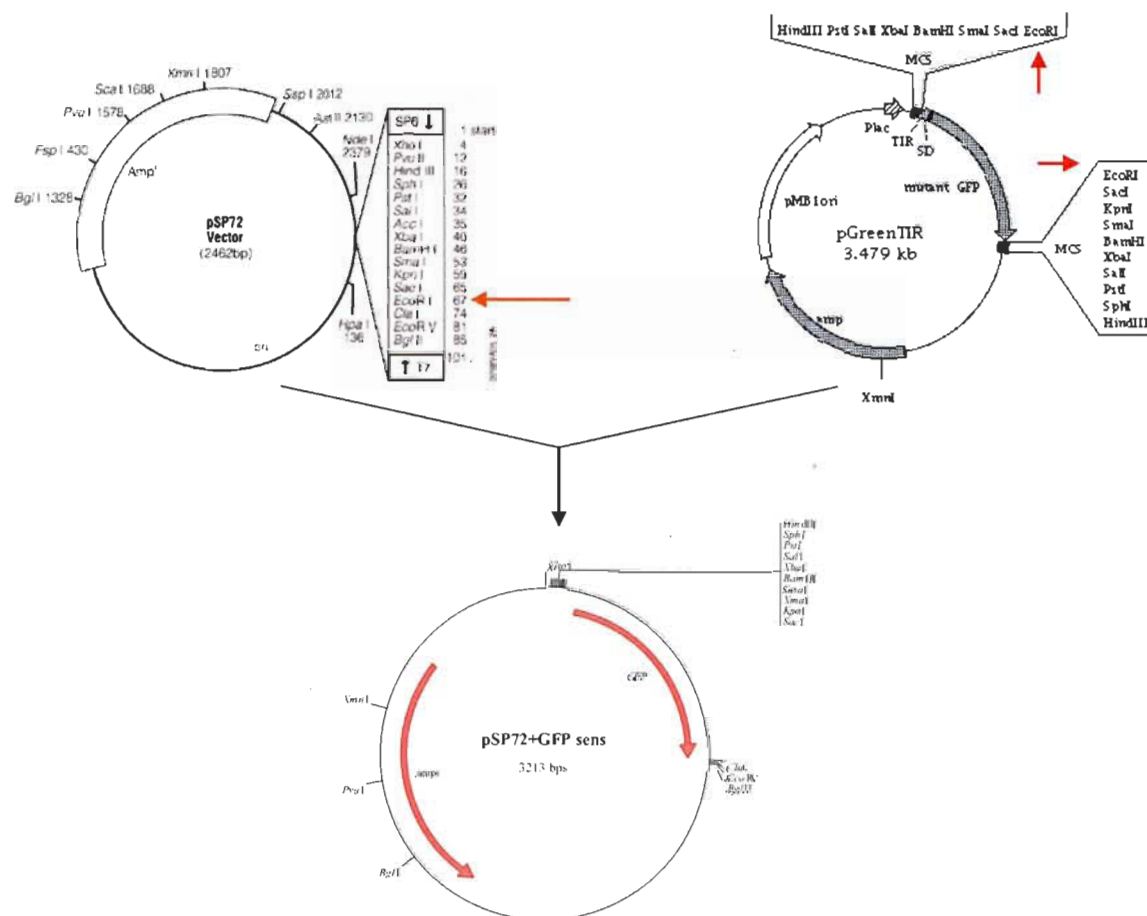


FIGURE 5.1: Plasmidic construction with restriction enzyme *Eco* RI. Red arrows represent enzymatic digestion sites on plasmids pSP72 and pGreenTIR.

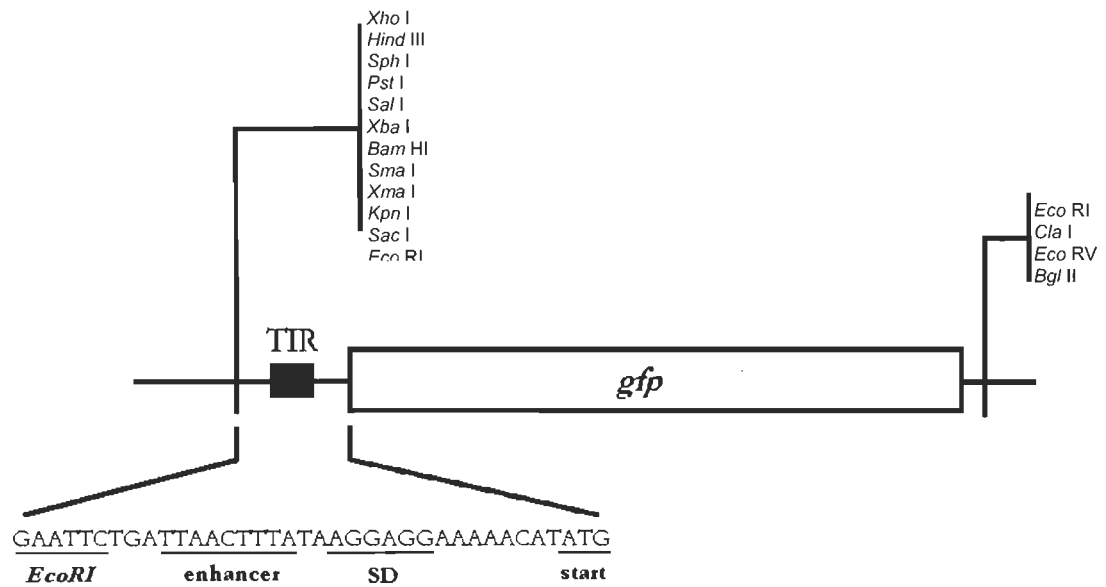


FIGURE 5.2: GFP cassette showing enzyme restriction sites and Translation Initiation Region (TIR): enhancer, SD (Shine-Dalgarno) sequence and star codon.

TABLE 5.1
Number of clones on LB amp100 and GFP positives clones
after electroporation of ligations

Ligations	GFP positive clones	Total clones	% GFP
Ligation 21-08-07	20	290	6.90
(pSP72-GFP + DNA ME digested <i>Bam</i> HI)	15	315	4.76
Ligation 24-09-07*	5	159	3.14
(pSP72-GFP + DNA ME digested <i>Sau</i> 3A)	3	164	1.83
Ligation 24-09-07*	3	126	2.38
(pSP72-GFP + DNA JB digested <i>Sau</i> 3A)	2	134	1.49
	41	440	9.32
	31	306	10.13
	28	415	6.75
Ligation 24-09-07	26	290	8.97
(pSP72-GFP + DNA JB digested <i>Sau</i> 3A)	33	65	9.04
	28	390	7.18
	24	265	9.06
	29	360	8.06
	47	510	9.22
	39	430	9.07
Ligation 24-09-07	36	390	9.23
(pSP72-GFP + DNA ME digested <i>Sau</i> 3A)	38	410	9.27
	29	315	9.21
	26	290	8.97
	36	340	10.59
	37	405	9.14
	47	480	9.79
	39	430	9.07
	54	500	10.80
Ligation 26-11-07	37	400	9.25
(pSP72-GFP + DNA DJ digested <i>Sau</i> 3A)	28	390	7.18
	49	478	10.25
	31	405	7.65
	38	510	7.45

	38	408	9.31
	52	495	10.51
	48	436	11.01
Ligation 26-11-07	50	420	11.90
(pSP72-GFP + DNA JB digested <i>Sau</i> 3A)	40	310	12.90
	39	400	9.75
	37	290	12.76
	46	305	15.08
	34	360	9.44
	44	407	10.81
Ligation 26-11-07	29	392	7.40
(pSP72-GFP + DNA ME digested <i>Sau</i> 3A)	31	405	7.65
	34	375	9.07
	44	400	11.00
	33	390	8.46
Ligation 26-11-07	40	460	8.70
(pSP72-GFP + DNA MS digested <i>Sau</i> 3A)			
Total	1538	16 955	

* 50µl were plated on LB amp100 media

TABLE 5.2
BLAST results for cloned sequences (GFP positives clones)
similar to bacterial DNA

Clones	Length (nt)	BLAST results in microbes genomic bank	Similarity (nt/nt)	Similarity %	Nucleotidic position
ME4-11-07	418	<i>Bacillus cereus</i>	158/171	92%	(249-418)
		<i>Eubacterium siraeum</i>	50/50	100%	(189-218)
ME8-11-07	596	<i>Bacillus cereus</i>	286/289	98%	(308-596)
		<i>Bacteroides capillosus</i>	81/93	87%	(209-297)
ME1-12-07	211	<i>Cyanothece sp.</i>	198/220	90%	(1-211)
ME16-12-07	403	<i>Cyanothece sp.</i>	305/331	92%	(77-403)
DJ5-12-07	757	<i>Escherichia coli</i>	757/758	99%	(1-757)
JB8-12-07	281	<i>Bacillus cereus</i>	144/156	92%	(125-279)
JB10-12-07	404	<i>Salmonella enterica</i>	321/404	79%	(1-390)

TABLE 5.3
BLAST results for cloned sequences (GFP positives clones) similar to human DNA

Clones	length (nt)	BLAST results in human genomic bank (position of the sequence)	Similarity (nt/nt)	Similarity %	Nucleotidic position
MS6-08-07	712	<i>Hs</i> chromosome 15	705/712	99%	(1-712)
ME7-08-07	167	<i>Hs</i> chromosome 12	167/167	100%	(1-167)
ME8-08-07	198	<i>Hs</i> chromosome 2	195/198	98%	(1-198)
2B-09-07	137	<i>Hs</i> chromosome 16	137/137	100%	(1-137)
6B-09-07	137	<i>Hs</i> chromosome 16	137/137	100%	(1-137)
7B-09-07	231	<i>Hs</i> chromosome 7	225/231	97%	(1-231)
3J-10-07	354	<i>Hs</i> chromosome 4	154/157	98%	(1-156)
		<i>Hs</i> chromosome 4	108/108	100%	(247-354)
		<i>Hs</i> chromosome 2	99/99	100%	(153-251)
6J-10-07	20	<i>Hs</i> chromosome 4	17/17	100%	(1-17)
2M-10-07	119	<i>Hs</i> chromosome 11	118/119	99%	(1-119)
JB3-11-07	479	<i>Hs</i> chromosome 6	354/354	100%	(126-479)
		<i>Hs</i> chromosome 3	130/130	100%	(1-130)
JB12-11-07	565	<i>Hs</i> chromosome 7	275/275	100%	(1-275)
		<i>Hs</i> chromosome 2	85/89	95%	(265-352)
		<i>Hs</i> chromosome 14	118/119	99%	(337-454)
JB14-11-07	426	<i>Hs</i> chromosome 9	162/162	100%	(265-426)
		<i>Hs</i> chromosome 2	153/153	100%	(1-153)
		<i>Hs</i> chromosome 16	119/119	100%	(150-268)
JB15-11-07	719	<i>Hs</i> chromosome 7	674/675	99%	(45-719)
JB16-11-07	454	<i>Hs</i> BAC clone	454/454	100%	(1-454)
JB17-11-07	245	<i>Hs</i> chromosome 4	245/245	100%	(1-245)
ME1-11-07	243	<i>Hs</i> BAC clone	174/174	100%	(70-243)
		<i>Hs</i> chromosome X	74/74	100%	(1-74)
ME4-11-07	418	<i>Hs</i> BAC clone	81/87	93%	(14-100)
ME6-11-07	348	<i>Hs</i> chromosome 1	279/279	100%	(70-348)
		<i>Hs</i> chromosome 8	73/73	100%	(1-73)
ME8-11-07	596	<i>Hs</i> chromosome 2	138/144	95%	(40-183)
ME11-11-07	418	<i>Hs</i> chromosome X	418/418	100%	(1-418)

ME15-11-07	449	<i>Hs</i> chromosome 11	375/375	100%	(75-449)
		<i>Hs</i> chromosome 21	79/79	100%	(1-79)
ME16-11-07	357	<i>Hs</i> 18S ribosomal RNA	278/278	100%	(80-357)
		<i>Hs</i> chromosome 4	84/84	100%	(1-84)
ME19-11-07	153	<i>Hs</i> chromosome 2	152/157	96%	(1-153)
ME22-11-07	465	<i>Hs</i> FOSMID clone	465/465	100%	(1-465)
ME23-11-07	463	<i>Hs</i> FOSMID clone	462/463	99%	(1-463)
ME3-12-07	46	<i>Hs</i> chromosome 2	46/46	100%	(1-46)
ME7-12-07	64	<i>Hs</i> chromosome 11	64/64	100%	(1-64)
ME12-12-07	208	<i>Hs</i> chromosome 19	67/67	100%	(1-67)
		<i>Hs</i> BAC clone	144/146	98%	(63-208)
ME13-12-07	546	<i>Hs</i> chromosome 15	372/376	98%	(1-376)
		<i>Hs</i> chromosome 15	157/177	88%	(372-546)
ME16-12-07	403	<i>Hs</i> chromosome 3	76/76	100%	(1-76)
ME19-12-07	312	<i>Hs</i> chromosome 3	215/215	100%	(3-217)
		<i>Hs</i> chromosome 14	98/99	98%	(214-312)
ME20-12-07	240	<i>Hs</i> chromosome X	237/340	98%	(1-240)
ME21-12-07	239	<i>Hs</i> chromosome X	239/239	100%	(1-239)
ME22-12-07	151	<i>Hs</i> chromosome 9	151/151	100%	(1-151)
ME24-12-07	387	<i>Hs</i> chromosome 20	127/127	100%	(261-387)
		<i>Hs</i> chromosome 3	264/266	99%	(1-266)
MS3-12-07	242	<i>Hs</i> chromosome 5	179/181	98%	(62-242)
		<i>Hs</i> chromosome 13	67/67	100%	(1-67)
MS7-12-07	253	<i>Hs</i> chromosome 7	126/132	95%	(122-253)
		<i>Hs</i> chromosome 4	122/126	96%	(1-126)
MS8-12-07	258	<i>Hs</i> chromosome 7	131/137	95%	(123-258)
		<i>Hs</i> chromosome 4	123/127	96%	(1-127)
MS9-12-07	115	<i>Hs</i> chromosome 13	113/115	98%	(113-115)

MS12-12-07	202	<i>Hs</i> chromosome 7	202/202	100%	(1-202)
MS18-12-07	419	<i>Hs</i> chromosome 22	412/420	98%	(1-419)
MS23-12-07	155	<i>Hs</i> chromosome 2	154/155	99%	(1-155)
MS24-12-07	112	<i>Hs</i> chromosome 12	112/112	100%	(1-112)
DJ1-12-07	172	<i>Hs</i> chromosome 2	172/172	100%	(1-172)
DJ4-12-07	487	<i>Hs</i> chromosome 18	487/487	100%	(1-487)
DJ7-12-07	178	<i>Hs</i> chromosome 10	95/95	100%	(1-95)
		<i>Hs</i> chromosome 6	93/95	97%	(85-178)
DJ8-12-07	767	<i>Hs</i> chromosome 4	767/767	100%	(1-767)
DJ12-12-07	101	<i>Hs</i> chromosome 2	101/101	100%	(1-101)
DJ18-12-07	187	<i>Hs</i> chromosome 5	187/187	100%	(1-187)
DJ22-12-07	64	<i>Hs</i> chromosome 2	61/64	95%	(1-64)
DJ6-12-07	214	<i>Hs</i> chromosome 10	76/76	100%	(1-76)
		<i>Hs</i> chromosome 19	142/142	100%	(73-214)
DJ19-12-07	206	<i>Hs</i> chromosome 22	88/88	100%	(75-162)
		<i>Hs</i> chromosome 10	58/61	95%	(147-206)
		<i>Hs</i> chromosome 10	44/44	100%	(36-79)
		<i>Hs</i> chromosome 4	40/40	100%	(1-40)
DJ23-12-07	344	<i>Hs</i> chromosome 5	344/344	100%	(1-344)
JB6-12-07	419	<i>Hs</i> chromosome 11	311/311	100%	(109-419)
		<i>Hs</i> chromosome 15	112/112	100%	(1-112)
JB7-12-07	300	<i>Hs</i> chromosome 14	118/125	94%	(177-300)
		<i>Hs</i> chromosome 3	167/183	91%	(1-187)
JB9-12-07	251	<i>Hs</i> chromosome 9	149/149	100%	(103-251)
		<i>Hs</i> chromosome 5	106/106	100%	(1-106)
JB11-12-07	74	<i>Hs</i> chromosome 17	74/74	100%	(1-74)
JB12-12-07	96	<i>Hs</i> chromosome 20	96/96	100%	(1-96)

JB13-12-07	69	<i>Hs</i> chromosome 6	61/64	95%	(1-64)
JB16-12-07	200	<i>Hs</i> chromosome 10	170/170	100%	(31-200)
		<i>Hs</i> BAC clone	35/35	100%	(1-35)
JB17-12-07	301	<i>Hs</i> chromosome 13	301/301	100%	(1-301)
JB18-12-07	157	<i>Hs</i> chromosome 5	157/157	100%	(1-157)
JB19-12-07	107	<i>Hs</i> chromosome 14	107/107	100%	(1-107)
JB23-12-07	786	<i>Hs</i> chromosome 2	638/639	99%	(97-730)
		<i>Hs</i> chromosome 5	96/96	100%	(1-96)
JB24-12-07	590	<i>Hs</i> chromosome 4	200/200	100%	(217-416)
		<i>Hs</i> chromosome 10	179/179	100%	(412-590)
		<i>Hs</i> chromosome 6	161/161	100%	(1-161)

TABLE 5.4
Predicted promoters with molquest application for all 69 cloned sequences

Clones	Inserts length	Number of predicted promoters	Promoter position
MS6-08-07	712	2	681 130
ME7-08-07	167	1	131
ME8-08-07	198	1	174
2B-09-07	137	1	108
6B-09-07	137	1	71
7B-09-07	231	1	128
3J-10-07	354	1	317
6J-10-07	20	0	
2M-10-07	119	1	96
JB3-11-07	479	1	342
JB12-11-07	525	1	334
JB14-11-07	426	1	340
JB15-11-07	719	2	567 116
JB16-11-07	454	2	380 61
JB17-11-07	245	1	216
ME1-11-07	243	1	79
ME4-11-07	418	1	207
ME6-11-07	348	1	98
ME8-11-07	596	2	545 84
ME11-11-07	418	1	164
ME15-11-07	449	2	430 101
ME16-11-07	357	1	69
ME19-11-07	153	1	104
ME22-11-07	465	1	151
ME23-11-07	465	1	151
ME1-12-07	211	1	161

ME3-12-07	46	0	
ME7-12-07	64	0	
ME12-12-07	208	1	158
ME13-12-07	546	2	493 129
ME16-12-07	403	1	270
ME19-12-071	312	1	112
ME20-12-07	240	1	115
ME21-12-07	239	1	115
ME22-12-07	151	1	68
ME24-12-07	387	1	121
MS3-12-07	242	1	73
MS7-12-07	253	1	192
MS8-12-07	258	1	193
MS9-12-07	115	1	62
MS12-12-07	202	1	149
MS18-12-07	419	1	202
MS23-12-07	155	1	67
MS24-12-07	112	0	
DJ1-12-07	172	1	138
DJ4-12-07	487	2	445 64
DJ5-12-07	757	2	171 481
DJ7-12-07	178	1	61
DJ8-12-07	767	2	703 168
DJ12-12-07	101	1	65
DJ18-12-07	187	1	144
DJ22-12-07	64	0	
DJ6-12-07	214	1	107
DJ19-12-07	206	1	156
DJ23-12-07	344	1	212
JB6-12-07	419	1	240
JB7-12-07	300	1	122
JB8-12-07	281	1	155
JB9-12-07	251	1	108
JB10-12-07	400	1	226

JB11-12-07	74	0	
JB12-12-07	96	1	73
JB13-12-07	69	0	
JB16-12-07	200	1	157
JB17-12-07	301	1	176
JB18-12-07	157	1	95
JB19-12-07	107	1	84
JB23-12-07	786	2	215 705
JB24-12-07	590	2	103 415

CHAPITRE 6

CONCLUSION GÉNÉRALE

La découverte de bactéries dans des endroits que l'on croyait stériles [12] et l'augmentation de bactériémies lors de transfusions sanguines [29] remettent en question le dogme de la stérilité du sang. La possibilité de découvrir une microflore dans le sang semble de plus en plus plausible puisque certaines recherches ont démontrés la présence d'ADN bactérien dans des plaquettes sanguines [24] et dans le sang veineux de sujets sains [44, 45]. Les bactériémies développées suite à une transfusion de sang pourraient hypothétiquement s'expliquer par une réponse immunitaire d'un receveur qui ne possède pas les mêmes espèces de bactéries que celles du donneur.

L'hypothèse émise stipule que le sang veineux d'un humain d'apparence saine contient des bactéries vivantes. Le premier objectif étant de mettre en évidence la présence d'ADN bactérien à l'aide de trois techniques différentes : PCR conventionnelle, PCR quantitative et trappe à promoteurs bactériens. Pour démontrer la viabilité, la PCR quantitative a été utilisée pour vérifier s'il y avait une croissance bactérienne suite à une incubation de 5 jours.

La première technique n'a pas permis la détection d'ADN bactérien et ceci est probablement dû à la faible sensibilité de cette technique puisqu'elle permet de détecter seulement une quantité d'ADN supérieure à 500 pg [48]. La quantité d'ADN bactérien dans le sang est peut-être inférieure à cette quantité. En se basant sur le fait que 5pg correspondent à approximativement 1000 cellules bactériennes, une amplification visible sur gel d'agarose correspondrait à 1 million de bactéries (si l'on considère qu'une bactérie possède une seule copie du gène codant pour l'ARNr 16S).

La seconde technique a permis d'obtenir quelques résultats attendus : présence d'ADN bactérien dans les échantillons (amplification avant le contrôle négatif) et croissance suite à une incubation. Par contre, l'absence de résultats reproductibles et la

positivité du contrôle négatif (parfois supérieur aux échantillons testés) ne nous permettent pas de prendre en considération les résultats. De plus, même si les anti-coagulants utilisés (héparine et citrate) ne sont pas supposés avoir d'effet sur la croissance, ceux-ci peuvent limiter la réaction PCR. Le SPS qui est couramment utilisé pour les hémocultures est reconnu pour inhiber la réaction PCR [52]. C'est pourquoi nous avons utilisé d'autres anti-coagulants, mais ils possèdent peut-être des éléments limitants également. La présence d'ADN bactérien et la viabilité n'ont pu être démontrés par cette technique.

La dernière technique, qui consistait à séquestrer des promoteurs bactériens dans un plasmide contenant la cassette GFP, a permis de détecter de l'ADN bactérien (promoteur bactériens) dans le sang de 3 sujets et ce, de manière aléatoire (sans détecter un gène précis comme pour l'ADNr16S). Les séquences insérées en amont de la cassette GFP sont similaires à six espèces bactériennes : *Bacillus cereus*, *Eubacterium siraeum*, *Bacteroides capillosus*, *Cyanothece sp.*, *Escherichia coli* et *Salmonella enterica*. Tous, sauf *Cyanothece sp.*, causent des bactériémies et/ou des septicémies chez l'humain [53-57]. *Cyanothece sp.* est une cyanobactérie [58] et elle n'est pas reconnue pour causer des infections chez l'homme. Les bactéries identifiées qui causent des bactériémies sont possiblement présentes normalement dans le sang de certains individus et, lors d'une transfusion ou lorsque le système immunitaire est affaibli, celles-ci causent des infections du sang. L'origine de ces bactéries n'a pu être déterminée puisqu'elles peuvent également provenir d'une contamination lors du prélèvement du sang.

En plus de mettre en évidence de l'ADN bactérien, cette technique a aussi démontré une interaction entre des séquences d'ADN d'origine humaine et la machinerie de transcription bactérienne. Découverte intéressante étant donné que cela n'avait jamais, à notre connaissance, été étudié auparavant. Les séquences d'ADN humaines contiennent les régions consensus retrouvées dans les promoteurs bactériens. En plus d'avoir un pourcentage AT élevé et d'être des séquences non codantes, celles-ci se retrouvent en moyenne à 190pb en amont du TSS. Théoriquement, un promoteur peut se situer jusqu'à environ 200pb en amont du TSS [39]. Tous ces éléments sont

probablement les raisons pour lesquels ces séquences humaines interagissent avec la polymérase d'ARN bactérienne. Pour vérifier l'interaction réelle entre ces séquences et la polymérase, des études de «foot printing» pourraient être effectuées. Ainsi, il serait possible de déterminer si les régions consensus interagissent vraiment avec la sous-unité sigma de la polymérase. De plus, comme contrôle, il serait intéressant de mesurer la fluorescence. Ainsi, il serait possible de comparer l'émission de fluorescence produite par des séquences consensus et de les comparer à des séquences non consensus. Comme les séquences consensus interagissent plus spécifiquement avec l'ARN polymérase, une plus grande quantité de protéines fluorescentes seraient produites comparativement aux séquences non consensus. Ces résultats supposent que les possibilités du génome humain sont sûrement supérieures à ce que l'on croit et que l'activité promotrice bactérienne n'est qu'un début dans les fonctions possibles de l'ADN non codant.

Bien que les recherches bactériennes aient démontré qu'il a bel et bien de l'ADN bactérien dans le sang humain, la viabilité n'a malheureusement pu être démontrée. Pour ce faire, la technique de la PCR quantitative pourrait être améliorée en effectuant en premier lieu une courbe standard pour le gène codant pour l'ARNr 16S. Ainsi, les résultats seraient quantitatifs. De plus, les erreurs dues à la présence d'ADN humain pourraient être éliminées par l'établissement d'une courbe standard avec un gène eucaryote constitutif comme la β -actine. Les échantillons de sang étudiés proviendraient de tubes sans coagulant pour diminuer le plus possible la présence d'inhibiteurs. Pour déterminer si les bactéries proviennent réellement du sang et non de la peau lors du prélèvement, des échantillons prélevés sur un même individu à des intervalles réguliers sur une période déterminée pourraient nous indiquer si les bactéries sont toujours présentes ou si elles sont différentes à chaque fois, indiquant ainsi une contamination. La trappe à promoteur bactérien a permis de détecter de l'ADN bactérien, mais la nécessité de séquencer plusieurs échantillons pour retrouver des séquences bactériennes font d'elle une technique plutôt laborieuse et non spécifique à l'ADN bactérien.

Pour conclure, cette étude n'a pas permis de démontrer la présence de bactéries vivantes dans le sang humain, mais elle ne l'a pas réfutée non plus. Des éléments nous

laissent croire que ce pourrait effectivement être le cas. Le cas échéant, ceci pourrait expliquer les problèmes reliés aux transfusions sanguines et si le sang contient des bactéries colonisant une microflore, celles-ci pourraient servir de vecteurs pour des médicaments; comme par exemple, l'insertion d'un gène codant pour l'insuline chez les personnes atteintes du diabète de type I. Il est certain que tout ceci implique la caractérisation des bactéries hypothétiquement présentes dans le sang pour une éventuelle mise en culture. Pour ce faire, l'identification à l'espèce permettrait de mieux identifier les caractéristiques de croissance des bactéries tels les besoins nutritifs et les conditions en oxygène nécessaires à la mise en culture. Néanmoins, la présence d'ADN bactérien dans le sang demeure une piste intéressante pour réfuter le dogme de la stérilité du sang.

BIBLIOGRAPHIE

1. Prescott, L.M., J.P. Harley, and D.A. Klein, *Microbiologie*. 2e éd. française. ed. 2003, Bruxelles: De Boeck. xxvi, 1137 p.
2. Zeldovich, K.B., I.N. Berezovsky, and E.I. Shakhnovich, *Protein and DNA sequence determinants of thermophilic adaptation*. PLoS Comput Biol, 2007. **3**(1): p. e5.
3. Campanaro, S., Vezzi, A., Vitulo, N., Lauro, F.M., D'Angelo, M., Simonato, F., Cestaro, A., Malacrida, G., Bertoloni, G., Vall, G. and Bartlett, D.H., *Laterally transferred elements and high pressure adaptation in Photobacterium profundum strains*. BMC Genomics, 2005. **6**: p. 122.
4. Pundak, S., H. Aloni, and H. Eisenberg, *Structure and activity of malate dehydrogenase from the extreme halophilic bacteria of the Dead Sea. 2. Inactivation, dissociation and unfolding at NaCl concentrations below 2 M. Salt, salt concentration and temperature dependence of enzyme stability*. Eur J Biochem, 1981. **118**(3): p. 471-7.
5. Kroes, I., P.W. Lepp, and D.A. Relman, *Bacterial diversity within the human subgingival crevice*. Proc Natl Acad Sci U S A, 1999. **96**(25): p. 14547-52.
6. Watson, R.O. and J.E. Galan, *Campylobacter jejuni survives within epithelial cells by avoiding delivery to lysosomes*. PLoS Pathog, 2008. **4**(1): p. e14.
7. Tlaskalova-Hogenova, H., Stepánková R., Hudcovic T., Tucková L., Cukrowska B., Lodinová-Zádníková R., Kozáková H., Rossmann P., Bártová J., Sokol D., Funda D.P., Borovská D., Reháková Z., Sinkora J., Hofman J., Drastich P. and Kokesová A., *Commensal bacteria (normal microflora), mucosal immunity and chronic inflammatory and autoimmune diseases*. Immunol Lett, 2004. **93**(2-3): p. 97-108.
8. Grice, E.A., Kong H.H., Renaud G., Young A.C., Wolfsberg T.G., Turner M.L. and Segre J.A., *A diversity profile of the human skin microbiota*. Genome Res, 2008. **18**(7): p. 1043-50.
9. Aas, J.A., Paster B.J., Stokes L.N., Olsen I. and Dewhirst F.E., *Defining the normal bacterial flora of the oral cavity*. J Clin Microbiol, 2005. **43**(11): p. 5721-32.
10. Kazor, C.E., Mitchell P.M., Lee A.M., Stokes L.N., Loesche W.J., Dewhirst F.E. and Paster B.J., *Diversity of bacterial populations on the tongue dorsa of patients with halitosis and healthy patients*. J Clin Microbiol, 2003. **41**(2): p. 558-63.

11. Alexandrou, T.J., Hariprasad S.M., Benevento J., Rubin M.P., Saidel M., Ksiazek S., Thompson K., Boonlayangoor S. and Mieler W.F., *Reduction of preoperative conjunctival bacterial flora with the use of mupirocin nasal ointment*. Trans Am Ophthalmol Soc, 2006. **104**: p. 196-201.
12. Bik, E.M., Eckburg P.B., Gill S.R., Nelson K.E., Purdom E.A., Francois F., Perez-Perez G., Blaser M.J. and Relman D.A., *Molecular analysis of the bacterial microbiota in the human stomach*. Proc Natl Acad Sci U S A, 2006. **103**(3): p. 732-7.
13. Hayashi, H., M. Sakamoto, and Y. Benno, *Fecal microbial diversity in a strict vegetarian as determined by molecular analysis and cultivation*. Microbiol Immunol, 2002. **46**(12): p. 819-31.
14. Zhou, X., Bent S.J., Schneider M.G., Davis C.C., Islam M.R. and Forney L.J., *Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods*. Microbiology, 2004. **150**(Pt 8): p. 2565-73.
15. Eckburg, P.B., Bik E.M., Bernstein C.N., Purdom E., Dethlefsen L., Sargent M., Gill S.R., Nelson K.E., and Relman D.A., *Diversity of the human intestinal microbial flora*. Science, 2005. **308**(5728): p. 1635-8.
16. Dymock, D., Weightman A.J., Scully C. and Wade W.G., *Molecular analysis of microflora associated with dentoalveolar abscesses*. J Clin Microbiol, 1996. **34**(3): p. 537-42.
17. Lyons, S.R., A.L. Griffen, and E.J. Leys, *Quantitative real-time PCR for Porphyromonas gingivalis and total bacteria*. J Clin Microbiol, 2000. **38**(6): p. 2362-5.
18. Bartelt, M.A., *Diagnostic bacteriology : a study guide*. 2000, Philadelphia: F.A. Davis. xv, 500 p., [12] p. de pl.
19. Lefrere, J.J., *[The BOTIA project ("Blood and Organ Transmissible Infectious Agents"): a European collection of blood samples and an observatory of agents transmitted by blood transfusion or organ transplantation]*. Transfus Clin Biol, 2005. **12**(2): p. 93-4.
20. Kaur, P. and S. Basu, *Transfusion-transmitted infections: existing and emerging pathogens*. J Postgrad Med, 2005. **51**(2): p. 146-51.
21. Bihl, F., Castelli D., Marincola F., Dodd R.Y. and Brander C., *Transfusion-transmitted infections*. J Transl Med, 2007. **5**: p. 25.
22. Hillyer, C.D., Josephson C.D., Blajchman M.A., Vostal J.G., Epstein J.S. and Goodman J.L., *Bacterial contamination of blood components: risks, strategies,*

- and regulation: joint ASH and AABB educational session in transfusion medicine. Hematology Am Soc Hematol Educ Program, 2003: p. 575-89.
23. Narayan, S., *Microbes and blood transfusion*. Indian J Med Microbiol, 2001. **19**(3): p. 119-26.
 24. Cirino, F., Webley W.C., West C., Croteau N.L., Andrzejewski C. Jr. and Stuart E.S., *Detection of Chlamydia in the peripheral blood cells of normal donors using in vitro culture, immunofluorescence microscopy and flow cytometry techniques*. BMC Infect Dis, 2006. **6**: p. 23.
 25. Guevara, R.E., Tormey M.P., Nguyen D.M. and Mascola L., *Listeria monocytogenes in platelets: a case report*. Transfusion, 2006. **46**(2): p. 305-9.
 26. Gosiewski, T., A. Kasprzyk, and M. Strus, *[Comparision of the sensitivity of detection of bacteria in human blood using classic culture methods and molecular techniques: PCR and FISH]*. Med Dosw Mikrobiol, 2005. **57**(3): p. 319-25.
 27. Petershofen, E.K., Fislage R., Faber R., Schmidt H., Roth W.K. and Seifried E., *Detection of nucleic acid sequences from bacterial species with molecular genetic methods*. Transfus Sci, 2000. **23**(1): p. 21-7.
 28. Christensen, J.E., J.A. Stencil, and K.D. Reed, *Rapid identification of bacteria from positive blood cultures by terminal restriction fragment length polymorphism profile analysis of the 16S rRNA gene*. J Clin Microbiol, 2003. **41**(8): p. 3790-800.
 29. Fenollar, F. and D. Raoult, *Molecular diagnosis of bloodstream infections caused by non-cultivable bacteria*. Int J Antimicrob Agents, 2007. **30** Suppl 1: p. S7-15.
 30. Peters, R.P., van Agtmael M.A., Danner S.A., Savelkoul P.H. and Vandenbroucke-Grauls C.M., *New developments in the diagnosis of bloodstream infections*. Lancet Infect Dis, 2004. **4**(12): p. 751-60.
 31. Turenne, C.Y., Witwicki E., Hoban D.J., Karlowsky J.A. and Kabani A.M., *Rapid identification of bacteria from positive blood cultures by fluorescence-based PCR-single-strand conformation polymorphism analysis of the 16S rRNA gene*. J Clin Microbiol, 2000. **38**(2): p. 513-20.
 32. Mohammadi, T., Reesink H.W., Vandenbroucke-Grauls C.M. and Savelkoul P.H., *Optimization of real-time PCR assay for rapid and sensitive detection of eubacterial 16S ribosomal DNA in platelet concentrates*. J Clin Microbiol, 2003. **41**(10): p. 4796-8.

33. Anthony, R.M., T.J. Brown, and G.L. French, *Rapid diagnosis of bacteremia by universal amplification of 23S ribosomal DNA followed by hybridization to an oligonucleotide array*. J Clin Microbiol, 2000. **38**(2): p. 781-8.
34. Kubista, M., Andrade J.M., Bengtsson M., Forootan A., Jonák J., Lind K., Sindelka R., Sjöback R., Sjögreen B., Strömbom L., Ståhlberg A. and Zoric N., *The real-time polymerase chain reaction*. Mol Aspects Med, 2006. **27**(2-3): p. 95-125.
35. Rantakokko-Jalava, K., Nikkari S., Jalava J., Eerola E., Skurnik M., Meurman O., Ruuskanen O., Alanen A., Kotilainen E., Toivanen P. and Kotilainen P., *Direct amplification of rRNA genes in diagnosis of bacterial infections*. J Clin Microbiol, 2000. **38**(1): p. 32-9.
36. Baker, G.C., J.J. Smith, and D.A. Cowan, *Review and re-analysis of domain-specific 16S primers*. J Microbiol Methods, 2003. **55**(3): p. 541-55.
37. Chakravorty, S., Helb D., Burday M., Connell N. and Alland D., *A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria*. J Microbiol Methods, 2007. **69**(2): p. 330-9.
38. Snyder, L. and W. Champness, *Molecular genetics of bacteria*. 1997, Washington, DC: ASM Press. xxii, 504 p.
39. Brosius, J., Dull T.J., Sleeter D.D. and Noller H.F., *Gene organization and primary structure of a ribosomal RNA operon from Escherichia coli*. J Mol Biol, 1981. **148**(2): p. 107-27.
40. Zhang, Z., R.C. Willson, and G.E. Fox, *Identification of characteristic oligonucleotides in the bacterial 16S ribosomal RNA sequence dataset*. Bioinformatics, 2002. **18**(2): p. 244-50.
41. Rediers, H., Rainey P.B., Vanderleyden J. and De Mot R., *Unraveling the secret lives of bacteria: use of in vivo expression technology and differential fluorescence induction promoter traps as tools for exploring niche-specific gene expression*. Microbiol Mol Biol Rev, 2005. **69**(2): p. 217-61.
42. Kanhere, A. and M. Bansal, *Structural properties of promoters: similarities and differences between prokaryotes and eukaryotes*. Nucleic Acids Res, 2005. **33**(10): p. 3165-75.
43. O'Neill, M.C., *Escherichia coli promoters. I. Consensus as it relates to spacing class, specificity, repeat substructure, and three-dimensional organization*. J Biol Chem, 1989. **264**(10): p. 5522-30.

44. Nikkari, S., McLaughlin I.J., Bi W., Dodge D.E. and Relman D.A., *Does blood of healthy subjects contain bacterial ribosomal DNA?* J Clin Microbiol, 2001. **39**(5): p. 1956-9.
45. McLaughlin, R.W., Vali H., Lau P.C., Palfree R.G., De Ciccio A., Sirois M., Ahmad D., Villemur R., Desrosiers M. and Canh E.C., *Are there naturally occurring pleomorphic bacteria in the blood of healthy humans?* J Clin Microbiol, 2002. **40**(12): p. 4771-5.
46. Mori, R., Wang Q., Danenberg K.D., Pinski J.K. and Danenberg P.V., *Both beta-actin and GAPDH are useful reference genes for normalization of quantitative RT-PCR in human FFPE tissue samples of prostate cancer.* Prostate, 2008.
47. Dakhama, A., Macek V., Hogg J.C. and Hegele R.G., *Amplification of human beta-actin gene by the reverse transcriptase-polymerase chain reaction: implications for assessment of RNA from formalin-fixed, paraffin-embedded material.* J Histochem Cytochem, 1996. **44**(10): p. 1205-7.
48. Sambrook, J. and D.W. Russell, *Molecular cloning : a laboratory manual*. 3rd ed. 2001, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. 3 v. (various pagings).
49. Alberts, B., *Biologie moléculaire de la cellule*. 4e éd. ed. 2004, Paris: Flammarion Médecine-Sciences. xxxiv, 1463, [85] p.
50. Rueckert, A. and H.W. Morgan, *Removal of contaminating DNA from polymerase chain reaction using ethidium monoazide.* J Microbiol Methods, 2007. **68**(3): p. 596-600.
51. Hughes, M.S., L.A. Beck, and R.A. Skuce, *Identification and elimination of DNA sequences in Taq DNA polymerase.* J Clin Microbiol, 1994. **32**(8): p. 2007-8.
52. Fredricks, D.N. and D.A. Relman, *Improved amplification of microbial DNA from blood cultures by removal of the PCR inhibitor sodium polyanetholesulfonate.* J Clin Microbiol, 1998. **36**(10): p. 2810-6.
53. Evreux, F., Delaporte B., Leret N., Buffet-Janvresse C. and Morel A., *[A case of fatal neonatal Bacillus cereus meningitis]*. Arch Pediatr, 2007. **14**(4): p. 365-8.
54. Yoshioka, H., Sakata H., Fujita K., Kakehashi H., Oka T. and Kaeriyama M., *Septicemia in immunocompromised children and their intestinal flora.* Acta Paediatr Jpn, 1990. **32**(6): p. 656-62.
55. Brook, I., *Clinical review: bacteremia caused by anaerobic bacteria in children.* Crit Care, 2002. **6**(3): p. 205-11.

56. Ruff, M.E., I.R. Friedland, and S.M. Hickey, *Escherichia coli* septicemia in nonperforated appendicitis. Arch Pediatr Adolesc Med, 1994. **148**(8): p. 853-5.
57. Janyapoon, K., Korbsrisate S., Thamapa H., Thongmin S., Kanjanahareutai S., Wongpredee N. and Sarasombath S., *Rapid detection of Salmonella enterica serovar Choleraesuis in blood cultures by a dot blot enzyme-linked immunosorbent assay*. Clin Diagn Lab Immunol, 2000. **7**(6): p. 977-9.
58. Stockel, J., et al., *Global transcriptomic analysis of Cyanothece 51142 reveals robust diurnal oscillation of central metabolic processes*. Proc Natl Acad Sci U S A, 2008. **105**(16): p. 6156-61.

ANNEXE A

```
1  CAGGTGATCC GCCCGCCTAG GCCTCCCAAG TTGCTGGGAT TACAGGTGTG
51  AGCCACCATG CCTGGCCTAT TTATTTATTT TGTAAGCCAG TCACATTTAG
101 CAGTAGGGGG TAATAGTACT TAATTGTCTC TTGTTCTTTG GACATGTTTT
151 GGAAAGAGTC TTTTACTCCT CCTAACATAT AAAATATACT TATTTATTAT
201 GTAGTATCTT GTAGTGTTTG GGGTTTCTTC CTGGTTGATT CTGGTAGCGC
251 ACATGAAGAT TCATTGAAGA TTCATGATGA ATAAAGAAGT ATGCTTTATT
301 CATCCTGATA TTGCTAAAT ATATTTAGTG CCTGGCAAAT AGGTGCTCAA
351 AAAAAAAAAA GAGGATTAGC AAGAGGATTA AATGATTGGT TCTTATAGAA
401 ACTTCTGATT TGTTAAATGA TAGATAATTA CACAATTACA GAATATTTTG
451 CTAGAGATAG ATTTTGAATA TAGATAGATT TTGGAGATGG ATTTTTGCCT
501 CTATGACAAT GTATTTTTTT AAAAATCACA GCACCTAGAT ATATAGATGC
551 AGACTATTTG GTAAAAATGA GCAGAACAAT GTTCTTCACA TTTCTCCAT
601 ATAAATGAGA GAAATACTTC ATATTAAGCA GGACTCTCTT ATTACAGGAA
651 TCACTGACCC TGGAGGATGT GGCTGTGGAG TTCACTTGGG AGGAGTGGCA
701 GCTCCTCAGC CCTGCTCAGA AGGACCTGTA CCGAGACGTG ATGTTGGAGA
751 ACTATAGCAA CCTCGTGTCA GTGGGTGAGG ACAGCTGCCC TGTGTCACTC
801 AGAGAGTACC CAGTCAAAGG CCTTTGCTTT CTCAGTTTCC AAAGCTCTGG
851 AGGGCCTGTG GTGCTCTGAA ATGTAGATTG CGGTACCCTC TCTGGCCCCA
901 GATAAAGAAT GGTAATGGGC CCCCTTTTAA AGCTCTGGAG GGCCTGTGGT
951 GCTCTGAAGT GGTAGATTCT GTACCCTCTC TGGCCCCAGA TAAAAGAGTG
1001 TAATGTGCCC CCTTTAGAGG GAAAATCCCT TTGCTTTGCA GACACATAAA
1051 TTACATAGTT CCTAAAACAG AACATGGTCC CATGTTGATA GACCACAGCT
1101 CAATTCAATG AGCCTAAGTT GTGTATCATT TCCTATGGAC AGGGTATCAA
1151 GCCAGCAAAC C
```

Séquence insérée dans le vecteur pDrive

ANNEXE B

MS6-08-07:

```

1   CTCTGTGTGG CCCCTCCAC TTTGTACTAG GGTTAGTCTG TATGACCGAT
51  GGATTATAGC CTATGTGATG GGATGTCACT TCCGAGATTA CTCCATGACA
101 CACTGTGGCT TCCACCATAA GCTGTCCCTT TATCTCTCCC TCATCACTCA
151 CTCTGAGGGA AGTCAACTGA TGCTTTACAA GGTGCCCAG GCAGCCAGGG
201 AAGAGGCTGT GGTGAGGCTC TGAGGTCTCC TGCTGACCGC CAGTGAACAA
251 TTGACGCCAC CACCAGCCCT GCGAGTGGGG ATGGAAGAAG GTCCTGCAGC
301 CCCAGCTAGA CTGTAGCCTC GAGAAGGACC CTGAGCCAGT TACACCACGC
351 TTGGATTTAT GAGCCTCAGA AATGGTGTAA GAAATATATG ATGGTTATCT
401 TAGGCTGCTA AATTTTGGGA TGATTTTGTG ATGCGGCAAT AGATGACTAG
451 TACACCACAA TATCAGACTT CGACTAATCA TTCCTGATTT AAAAAAAAAA
501 AAAAAAGAACT GCAATGGAAT GACATACAAA TGCACCAGAA TGGCTAAAGT
551 TAAAAAGACT GACAACACTA AACATCAAAG ATGATGGACC AAGCAGGACT
601 CTCGTATGTT GTTACAGTGG GAGTATAAAA CAATATGAAC ACAATTGAAA
651 AGGTCTAGCA GTTTCGTGTA TACTCAACAT ACATCTGCCT ATGACCTACC
701 TACATCTACC TG

```

Position of predicted promoter: **87** bases upstream the gfp cassette

ME7-08-07:

```

1   CAAGCTCACT CTAAAATAGA AAGGATAGAG AAGTACATAA ATCATAGCAC
51  AAGTATGCAT CTACAAAAGG CTACGGGAAC AGAGAAGGAA ATAATGAACT
101 TTGTGGAAGC AGAGATGTTA TAAAGATGAC AATGCAGCAG AGCATCAGGG
151 AATCATCAGG TGGACAA

```

Position of predicted promoter: **92** bases upstream the gfp cassette

ME8-08-07 :

```

1   AAGGGCAGAA TTCTGTTGGT ACCCACTGGA TGCTCACCTC CCCCAAGCCC
51  AAGGCTATCA TGAGTCAAGA GCAAAAGGGC TTCCAAGCTC ACGGCCCCCA
101 CTGAGGAGGT CTTGGGTTCA TGTTATCAGC ACTGAACTTT CTCTAAACAT
151 TTCAGGGAAC ATATATTTCC CACTTTGAAC TACACTTACC AGGGGGAT

```

Position of predicted promoter: **80** bases upstream the gfp cassette

2B-09-07:

```

1   AAATAATGAA ATCTCAATAG CTAACAAACA GTAAGCACGT GCTGATGCCA
51  GTAGCTCTGC TAAGAGCATG ATATATCAGC TCCTCAGAAC AACTCTATGA
101 TACAGATAAA CTATTGCAA ATCACATACC TGATACA

```

Position of predicted promoter: **85** bases upstream the gfp cassette

6B-09-07:

```

1   CTACCCATCT GGACTAAATA ATGTCAATAA ATAATGGAAT CTAACCTCTCT
51  CCCCATTGGAA GAATTTTGTAG CCAGAGGGAG AAATTGTAGC ATTTCAGGCC
101 TCTGAAGTCA GGTGTTGGAC CTGGAACTA GCAAAAT

```

Position of predicted promoter: **122** bases upstream the gfp cassette

7B-09-07:

```

1   CTTTTCATTT TACTGAGGAA AATGTAAACA TTCCATCTGA CATCATCTCC
51  CTCAATTCCT GCCTCTCCAT TTGTCCTGTT CCTCCTGTCA CAGTGCAAGG
101 GATGTTTGTC AGTGACAAAG TTCCAGGTTG TGAACACCTC AAAGCCTGAC
151 TTAATGTGGC TTTTTCCTGG AAGTCCAGGT GAAGGGCAGG CTTCAGGGCT
201 GGTCCATCCT GCGGCTCATT GACTGCCTCA A

```

Position of predicted promoter: **159** bases upstream the gfp cassette

3J-10-07 :

```

1   CATCATAACC GTCTTCATCC CCCTCATCTT TACATGGAGT ACGCTAAAGG
51  GGAAGAGAAA GAATAACATT GGTTTGTCTG TCTTCGGTGG CAAAGGTGAA
101 AGAAAATTCA CATATAAGTG GATACACACA GTTGAAACCA ATGTTGTTCA
151 AGGATCTGCA TTGTAGGGAT TACTGCCTTA TATAGTCCAA AGGTCTAAGG
201 AAGCACCAGG GATGCATTCA CACTTATGTA CACACTGATG TCATCAGATC
251 TGAACCTGGG AATTAAAATG ATGCTAATTT TCCAACAAAA GCATGATTTA
301 AATTTCACCA TTTTTTTCAC TGCTTTTTTT TTTTCTGTTG CAGGATTCTC
351 TAAG

```

Position of predicted promoter: **99** bases upstream the gfp cassette

6J-10-07 :

```

1   ACCTAAGCTC TTGAGTTCAA

```

2M-10-07 :

```

1   AGTATGTGGC CTACCTTGTT TCCATACATA CTGTGCTCCT CCCAGACTAA
51  ATTATTTAGA AGCAAATGTC AGCTATCACA TGGTTTAATC TGTAAGTACT
101 TAAGATGTGT CCCCAAGAG

```

Position of predicted promoter: **79** bases upstream the gfp cassette

JB3-11-07 :

```

1  ATATTACTGT GTTCTTGGTC CTAGTTTCTC CTCGGCATGG AAGGAAGCTG
51  GTAAAAGGTT AGAAATGCAG GCAATGGCTG CAAGTCGAGA TTGCAGTTGA
101 CTGTTACATC AGGGAACAGG CATTATGATC ACCATTCTAC TCTCTACTTC
151 TATGAAATCA ACTTTTTTAG ATTCCACATA TGAGTGAGAT CATGAGGTAT
201 TTGTCTTTTT CCGCCTGGCT TATTTCAATT AACACAATGT CTTCCAGGTT
251 TATCCATGTT GTCACAAATT ACAGGATTTT CTTCTTTTAA TAGCAGGACA
301 GTATTCCATT GTGTATATAT ATCACATTTT TTAATCCATT CATCCACTGA
351 TGGACACTTA GGTGATTCT CTGTCTTGGT TATTGTGAAT TCTGCTGCAG
401 TGAACACGGA TGTTAGATAT CTCTTCGACA TATTGATTTT ACTTATTTTC
451 ATATTTTTTC ATTACCTTAC AAAATACTT

```

Position of predicted promoter: **193** bases upstream the gfp cassette

JB12-11-07 :

```

1  TTTCCCCATT GCAAAGGAAC AACACCGTTA CCTCTCTGGC TACTCAAGTC
51  AGAAACCTGG AAGTCGTCTT TCATTCCCCT CCCCTCCTTT ACCCACTGTA
101 TGAGCAAATC CTATTGATTT AATTTCCAAA TTTGTGCAAG AAACATATCTT
151 TCACTTAGAG TACACTGACA TCACCTTTTA CCTGACCATT ATCATAATAT
201 CATAATTCTT TTTCTATCTC TTCTTTTGCT TCATTTTCAC CCAATTCATC
251 TTCAAACAAC AGTGAGAGTG ATCTTCATTT CATCATATAA GACATATTAC
301 AAAGCCCAA TAAATAAAAC TGATAATGGT GTGGAGAGAT CAGCCACTGT
351 GTGGCCCCCA CAGAGCCACA GGCAACACAG GTTGGCGAAC AGAAAAGTGA
401 TGAGCCCTCT TTTAGCTCCA AATCAAAAGG AGCATGATGG GCCAGACTTG
451 GATCTTTGCC TGTAGTATAT ACATTTTTTG AGACAGGGTC TCGCTCTGTT
501 GCCCAGGTTG GAGTTGCAGT GGTGCCATCT CAGCTCACTG CAACCTCCAG
551 CTCCTGGGCT TGAGT

```

Position of predicted promoter: **287** bases upstream the gfp cassette

JB14-11-07:

```

1  ATATCATTTG CCAACAAGGA CAATTTTACT TCTTTTATTT TCCAATTTGG
51  ATTACCTTTA TAGCCTTCTC TTGTGTGACT GCTCTAGCTA GAACCTTCCG
101 TACTATGTTG CATAACAGTG GTGAAAATGG GCATCTTTGT TGTGTTCCAG
151 ATCAGAATGT CCCAAGCTTG GGACAGTCTG GGAAGGCAGT GGCCATCCCA
201 TCCAGATGAG TACATCCCTC TCTCCTTGCC TACTTCCCTC CTACCAGCCG
251 TCGCGGAGGC CACTGATCGA GACAAGAAGA CAACAGGATG TACTTTCTTC
301 TTGTCAAGAT AGAATTTCAT GCCTGGCTAA TCAACTAGAA AGTCTGATTT
351 TTGCATGCCT CAGGAAAAAG CAATCATAGG TGGTTTTGAG AGGAGGGTTC
401 TCTAGTGTTT CCTTCCCTCA GAGCCG

```

Position of predicted promoter: **141** bases upstream the gfp cassette

JB15-11-07:

```

1   AGAATAGAGA AGGCAGAAAT AATGCAGCAT ACCTGCAACC AATGGATCTT
51  GTGAGAACTC ACTATCAGGA GAACAGCATG GGGGATTATA ATTACATAGG
101 GATTATAATT CAAGATGAGA TTTGGTGGGG ACACAAAGCC TAGTCATATC
151 GTGGACCATG ATAATCTCAG AGGCTTTGGG AAGTGTGCTA ATATCTACAT
201 TTCCTGTAAT AACTTTTATA TCTAATCGGC ATTTAAAAAG CTGTTTTAAC
251 AACATTTGAG TGA CTGTCCT GATAGGCTCT GAAGGAATGC AAAGATTTGT
301 CTCTGTCATC CATCCATGGT CTCTGTCATC CATCCATGGT CTCTATCATT
351 TATCCATTCA TACAGAAGGT ATTTATGGAG ATACCTGAGT GCCCACTATG
401 CCCTAGGTAC TGTGTTTTTG AATTGACAGT TGAAAACATG GCAGTACTCT
451 GTTCAGGAAG TTTACTAACT GAACAGACCT GTGAGAAGCA GCGGGACTTT
501 TCCTGGGGGC AGTGGGTGGG CTTCAGGGGA ATCTTTGAAC CCTTTTAAAT
551 TATATATAAT ATTTTTTGTG TGTAAGCCCA TTTTAGGGGG AAGAAAGTTC
601 ATAGCTCTTA GTTTCTCAAG GAGTTTGGGT CCCTTAGACT AGTGGAAGA
651 TAAGGAGTTT AGGAGCTATA CAGTCATTAA TTTCTATCCT GGCCCTGCTG
701 TGATTCTGAG CAGATTATC

```

Position of predicted promoter: **208** bases upstream the gfp cassette

JB16-11-07:

```

1   TTTTTCGAAT TTTCTTTTAG CTCATCAGCT ATGGTTAGTG TTAGTATATA
51  TTACATGTGG CCGAAGACAA TTCTCCTTCT TCTTCTTCTT CCGGTGTGGC
101 CCAGGGAAGC CAAAAGATTG GACACCCCTG ATTTAGAGGA AAACAGAGGA
151 GGAAAATCCA GCAAAGGAGC AACCAACAAG AGGCAGGGAG AAATCTAGGA
201 AAGTATGCTC TAAGAGCCAA GAGAAGAAAG TGCTAAGGAG GGCAGTTAA
251 GAAGGAAGGA AGTGGTCAAT GAGACAAGCG TAGTTGGAAT AAGGAGGAAG
301 ATGAGTACCC AGAAGAGGTC ATTGGATTTG GTAATGTGGA AATCATGGAT
351 GACCTTTAAA AGGA AACTAA ACTCTAATTA AGGTGGGCTG GGAAAACACC
401 TGAGGAATGT AGAAAACAAT TTTAGACAAC TCTTCTGCAA AGTTTTACTA
451 TGAA

```

Position of predicted promoter: **130** bases upstream the gfp cassette

JB17-11-07 :

```

1   ATCAAATATT AATAAAGGAA TTGTAGTCAA AGTAGTTAAG CACATACAGG
51  CAGTTAGTGG AGAAGCTGTG TTTAGAACCC AAGCCAGCTG TTCTTCTTC
101 CCCAGGAGA GTGGAATTTT ACTACACCTT CCTCTTCCTA AACATGGCAC
151 AGATATTACA TTAAAGAAAT GGGGAAATCT ATGATTCTTT ATTTAGTGAG
201 TTCTAAACTA TGTATCTAAA ATGACATAAA ATGTGAATAT ACATG

```

Position of predicted promoter: **85** bases upstream the gfp cassette

ME1-11-07:

```

1   ATTAGCACAA TAAGGAAAGG ATATTCCCAT CAGTAAGTAA GCAGTGATAT
51  TTGGCCAAAT CTAGGTCATG ATCACGAGGT CAGGAGTTTG AGACCAGCCT
101 GACCAACATG GTGAAACCCT GTCTCTACTA AAAATACAAA AATTAGCCAG
151 GTGTGGTGGT ACGCACCTGT AATCCCAGCT ACTCAGGAGA CTGAGGCAGG
201 AGAATCACTT GAACCCAGGG GCGGAGGTT GAAGTGAGCC AAG

```

Position of predicted promoter: **220** bases upstream the gfp cassette

ME4-11-07 :

```

1   TACTCTATGT CCCCTACTGA AAACCTCAATC TAAGTTAAAA GGCTCATTTT
51  AAAACTAACG GGATGGAGAA AAATATACCA TGTTACACCA GCCAAATGAA
101 GCTCTAGCGT AAAAGGACAA ATGCTCTGTA GTAAACCTGC ACGCATGCAA
151 GCTACAATAG CTCCAGTTGG AGAACGCCAC CTACATCCTA TGTGTATGAT
201 ACATAAGGTT ATGTATTAAT TGTAGCCGCG TTCTAACGAC ACTATGTCCA
251 TATGGTGCAC TCTCACTACA ATCTGCTCTG ATGCCGCACA GTTAAGCCAG
301 CCCCGACACC CGCCAACACC CGCTGACGCG CCCTGACACC GTTGCCCGCT
351 CCCGGAATCG GCTTACAGAC AAGCTGAGAC CGACTCCGGG AGCTGCATGT
401 GTCAGAGGTT TTCAACGT

```

Position of predicted promoter: **267** bases upstream the gfp cassette

ME6-11-07 :

```

1   TCTCAACTCA GCTTCCTGAG TAGCTGAGAT TAAGCCATTG TACTCTGCTT
51  GTTATGAATT TTATATGTTG ATCAAACCCA GAAGGCAAAT TTCATTAAC
101 TAACAACATA AGGCATCTGT TACGGCTGGG TGCAGTGGCA CACTCCTGTA
151 ATCCCAGCAC TTTGGGAGGT CGAGGTGGGT AGATTACTTG AGGTCAGGAG
201 TTTGAGACCA GCCTGGCCAA CATGGTGAAA CCCCATCTCT ACCAAAAATA
251 CAAAAATTAG CCAGGCATGG TGGCGTAATC CCAGCTACTC AGGAGGCTGA
301 GACATGAGGA TTGCTTGAAC CCAGGAGGCA GGAGTTGCAG TGGGCTGA

```

Position of predicted promoter: **306** bases upstream the gfp cassette

ME8-11-07 :

```

1   CCGCCTTGAC CTCCCAAAGT GCCGGGATCA TAAAAAAGAA AATAACCCTT
51  GGCAGTCCCA AGCCTAGAAC TTAATATACC TAGCAGTGCC TGCTTTCACT
101 CACCTGAAAG TTAACCTACCA TGAAAAGAAA TCTACACAAT TCACATGGAA
151 ACAGTGGGGC CACCTGGAGG AAAATCAAGA TCCTCTATAG TCTACCTGCA
201 CGCATGCAAG CTTCAGCTGC TCAAGTTCTA TAGTGTCACC TAGATCGTAT
251 GTGTATGATA CATAAGGTTA TGTATTAATT GTATCCGCGT TCTAACGACA
301 ATATGTCCAT ATGGTGCACT CTCAGTACAA TCTGCTCTGA TGCCGCATAG
351 TTGAGCCAGC CCCGACACCC GCCAACACCC GCTGACGCGC CCTGACAGGC
401 TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA
451 GCTGCATGTG TCAGAGGTTT ACACCGTCAT CACCGAAACG CGCGAGACGA
501 AAGGGCCTCG TGATACGCCT ATTTTATAG GTTAATGTCA TGATAATAAT
551 GGTTCCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG CGCGGA

```


Position of predicted promoter: **107** bases upstream the gfp cassette

ME11-11-07:

```

1   TCAGGTGATC CACCCACCTC GGCCTCCCAA AGTGCTGGGA TTACAGGTGT
51  GAGTCACCAT GCCCGGCTGC TCCTACTATG CTTTCTGAA CTTTGGTTTT
101 TTTTCACTTA GACAAAAGTA TCCCAGACAT TCTTACATAT GAGTCAATTT
151 TTATCATTAAC AGTTCCAGAG TATGGATGTG GCAATTTATT TAACTATTCT
201 GTAGTCCTCC TCATCCATGT GAAGTTTGTT TACCAATCAG GGGTCATTTT
251 TATCATAAGC AATGCCACCT GGTAGTATAG CTATCATTCT ACCATTATTG
301 GATTCTCAGG AACATGGGAC ACTTAGGGCT TTCTCCTTCT ACTTTTGTCT
351 AGTATGAATA GAACTGCTTT GGAAATCTTT GTGGACATAT AAATATTTCA
401 CTAGGATGGA GTCCTATC

```

Position of predicted promoter: **309** bases upstream the gfp cassette

ME15-11-07 :

```

1   TTGTGGCCCC TGTGGACTCA GGGCCCATGT GTGCAGGGAG GAGGCACTCT
51  CCCCACGTGC TCCATTGGCCA GGGCTGATCA TGGGGTGTCA GAACCCACAC
101 TCGACCGATT TTAGGGTCAT TCCCCACAAA AGTCACCAGG CTGAGACCGA
151 GACAGTGCTT TCCCCCAATT TACTAAGTTG CAGAAGGTCA GACTCTGAAT
201 TCAAACCCGT GTCTAGCTCC CATCTCACCC AACTGAACCA TGCTCTGTTT
251 AGTTTCCGAT TTCCTCATAT AGTCGGCCCT CAGCATCCAT GGAGGATTGA
301 TTCCAGGGCC CCTGTGGACA CAAAATCGGT CAGTGTTCAA GTTCTTGACA
351 TAAAAATTGTG TGGCACTTCC ATGCAACCTA TGCACATCCT CCCATATGCT
401 TTTTTTTTTT TTTCTGAGA TAGTTTACT CTGTCACCTA GGCTGGCGT

```

Position of predicted promoter: **74** bases upstream the gfp cassette

ME16-11-07 :

```

1   ATCCTATAGC TAGTCCTGGA AAAGTACTGT GATTCAAGTG GAAAAGTACT
51  CTGATTTACT CTTTGGAATA CAATTTCTGG GATCCTCGTT AAAGGATTTA
101 AAGTGGACTC ATTCCAATTA CAGGGCCTCG AAAGAGTCCT GTATTGTTAT
151 TTTTCGTCAC TACCTCCCCG GGTCGGGAGT GGGTAATTTG CGCGCCTGCT
201 GCCTTCCTTG GATGTGGTAG CCGTTTCTCA GGCTCCCTCT CCGGAATCGA
251 ACCCTGATTC CCCGTCACCC GTGGTCACCA TGGTAGGCAC GGCGACTACC
301 ATCGAAAGTT GATAGGGCAG ACGTTCGAAT GGGTCGTCGC CGCCACGGGG
351 GGCGTGC

```

Position of predicted promoter: **344** bases upstream the gfp cassette

ME19-11-07 :

```

1   GCCTACTGTT TTACACAGTG GTTTTCCTTT TATTTTGCAT TCTTTTGAAA
51  ATGTTTTTCAA GGTGTCTTCA AATTCAGCAA AATCTAAATA CTATAAAAAA
101 TAAATAAATTC AATTCAACAG CAAAACAAAT AATCCAATTA AAAACAGGCC
151 AAA

```

Position of predicted promoter: **105** bases upstream the gfp cassette

ME22-11-07 :

```

1   AACGAGACAG AAAGTCAACA AGGATACCCA GGAATTGAAC TCAGCTCTGC
51  ACCAAGCAGA CCTAATAGAC ATCTACAGAA CTCTCCACCC CAAATCAACA
101 GAATATACAT TTTTTCAGC ACCACTCCAC ACCTATTCCA AAATTGACCA
151 CATAGTTGGA AGTAAAGCTC TCCTCAGCAA ATGTAAAAGA ACAGAAATTA
201 TAACAAACTA TCTCTCAGAC CACAGTGCAA TCAAAC TAGA ACTCAGGATT
251 AAGAACTCTA CTCAAAGCCG CTCAACTACA TGGAAACTGA ACAACCTGCT
301 CCTGAATGAC TACTGGGTAC ATAACGAAAT GAAGGCAGAA ATAAAGATGT
351 TCTTTGAAAC CAACGAGAAC AAAGACACCA CATACCAGAA TCTCTGGGAC
401 GCATTCAAAG CAGTGTGTAG AGGGAAATTT ATAGCACTAA ATGCCTACAA
451 GAGAAAGCAG GAAAG

```

Position of predicted promoter: **370** bases upstream the gfp cassette

ME23-11-07 :

```

1   AACGAGACAG AAAGTCAACA AGGATACCCA GGAATTGAAC TCAGCTCTGC
51  ACCAAGCAGA CCTAATAGAC ATCTACAGAA CTCTCCACCC CAAATCAACA
101 GAATATACAT TTTTTCAGC ACCACTCCAC ACCTATTCCA AAATTGACCA
151 CATAGTTGGA AGTAAAGCTC TCCTCAGCAA ATGTAAAAGA ACAGAAATTA
201 TAACAAACTA TCTCTCAGAC CACAGTGCAA TCAAAC TAGA ACTCAGGATT
251 AAGAACTCTA CTCAAAGCCG CTCAACTACA TGGAAACTGA ACAACCTGCT
301 CCTGAATGAC TACTGGGTAC ATAACGAAAT GAAGGCAGAA ATAAAGATGT
351 TCTTTGAAAC CAACGAGAAC AAAGACACCA CATACCAGAA TCTCTGGGAC
401 GCATTCAAAG CAGTGTGTAG AGGGAAATTT ATAGCACTAA ATGCCTACAA
451 GAGAAAGCAG GAA

```

Position of predicted promoter: **368** bases upstream the gfp cassette

ME1-12-07 :

```

1   TTTTTTATGG CTGCATAGTA TTCCATGGTG TATATGTGCC ACATTTTCTT
51  TATCCAGTCT ATCATTGATG GGCATTTGGA TTGGTTCCAA GTCTTTGCTA
101 TTGTGAATAG TGCCACAGTA AACATTACATA TGCATGTGTA GAACAATTTA
151 TAATCCTTTG GGTATATGCC CAGTAATGGG ATTGCTGGGT CAAATGGTAT
201 TTCTAGTTCT A

```

Position of predicted promoter: **106** bases upstream the gfp cassette

ME3-12-07 :

1 AATGTTT TAGT ATATATTACA TAGAAAACT ATTCCTAGA AAAAAT

ME7-12-07:

1 AGTCAGTTTA ATCCAACAGC GATTCTAATA TATTAGCTGC ATACAATGTT
51 TTTGTCTCAT AAAA

ME12-12-07:

1 TCCTACCCCG AGTCCTCCTC CTGGTCCTGA GTTCCAGGGT GGGCTTTGTT
51 AGGGACAAGA AGGATCCCAC GGAAAAACAA ACTACCATCA GAGAATACTA
101 TAAATACCTC TATGCAAATA AACTAGAAAA TCTAGAAGAA ATGGATAAAT
151 TCCTGGACAC ATACACCCAC CCAAGTCTAA ACCAGGAAGA AGTCAAATCC
201 CTGAGTAG

Position of predicted promoter: **106** bases upstream the gfp cassette

ME13-12-07:

1 ATCTGAGCCC AGGAGTTAGA GGCTGCAGTG AGCAAGGATC ATTCTACTGC
51 ACTCCAGCCT GGATAGCTGA GCAAGACCCT GTCTGAAAA TAATAAAATA
101 CAATACAATT AAGTTAAATA ATAAAAAATT AAAAAAAAAA TCCTGCGACA
151 TGAAATGACA CAGATGAACC TTGAAGCCAT TAAGCAAAGT GAAGTAAGCC
201 AGTCACAGGA CAAATACTAC ATGATTCCAC TGCAACGAGG TATCTAAAAT
251 AGGCAAACCTC ATACAAGCAA AGAGGGGCAC AGTGGTTGCC AGGGTAGGGG
301 CTTTCAGGAA ATGGGGAGTT GTTAATCAAA AAGTATAAAG TTTTCAAGTAAT
351 GCAAAATGAA CGCATTCTAG AGATCTCAGC TCACCGCAAC CTCCACCTCC
401 CAGGTTCAAG CGATTCTCCT GCCTCANCCT CGTGAGTAGC TGGGGTTACA
451 GGCATGTGCC ACCATGCCCG GCTAATTTTG TATTTT TAGT AGAGATGGGG
501 TTTCTCTATG TTGGTCAGGC TGGTCTCGAA CTCCTGACCT CAAGTG

Position of predicted promoter: **109** bases upstream the gfp cassette

ME16-12-07:

1 AAAACTGGCC TCTGCTTCTT TTTGTTTTTT AAAGAGACAG GGTCTCACTC
51 TGTCTGCTGG GCTAGAGTGC AGTGGCATGA TCGCCATTCT AACTGGTGTG
101 AGATGGTATC TCATTGTGGT TTTGATTTGC ATTTCTCTGA TGGCCAGTGA
151 TGATGAGCAT TTTTTCATGT GTCTGTTGGC TGCATAAATG TCCTTTTGAG
201 AAATGTCTGT TCATATCCTT TGCCCATTTT TTGATGGGTT GTTTGTTTTT
251 TTTCTGTGTA ATTTGTTTGA GTTCATTGTC AATTCTGGAT ATTAACCTT
301 TGTCAGAAGA GTAAATTGCA AAAATTTTCT CCCATTCTGT AGGTTGCCTG
351 TTCACCTCTGA CGGTAGTTTC TTTTGCGGGG CAGAACCTCT TTATTTTAAT
401 TAG

Position of predicted promoter: **189** bases upstream the gfp cassette

ME19-12-07:

```

1   AGACCATATC AATATGGAAA AACAAACAAA AAGCCTTAAT AGCCAATGCA
51  ATCCTAAGCA AGAATAAAGC CTGAGGCATT ACACTATACA ACTTCAGACT
101 ACACTATAAG GCTACAGTAA GCAAACAGCA TGGTACTGGT ACAAATCAG
151 ACACACAGAC CAAAGGAACA GAATAGAAAA CTCAGAAATA AAGCCACACA
201 ACTGTAAGAA TCTGATCCAG GAAGCCCAAC AGTTGGGTGG GCCACATGCT
251 TGCATGAAAT GCAAGGTCCC ACCCTGGCTA CTGGGGAGGA GAAGAAAGGC
301 AGGGTGGGGC AG

```

Position of predicted promoter: **266** bases upstream the gfp cassette

ME20-12-07:

```

1   ATTAGTATAT TAAAGTTTCA GAGAAATTGC GGTGGTAAAA TTCGTATTTA
51  ATATTGTTTA ACCCAGACTC TTACAAACTT GTTTGGTGGA AGAAATATTT
101 TTTATTAAATA AACTTCCGTC TACATTAGTG TTTCAACTAA ATACATCCTA
151 CAAAACTTTG AAGATAACGG GAAATTAAAA TAGTTGTAGG GGAAATGGAA
201 AGGAAGATAA AATCTGGGAT ATAATTTGAA GTTTTCAGTG

```

Position of predicted promoter: **181** bases upstream the gfp cassette

ME21-12-07:

```

1   ATTAGTATAT TAAAGTTTCA GAGAAATTCC AGTGGTAAAA TTCTTATTTA
51  ATATTGTTTA ACCCAGACTC TTACAAACTT GTTTGGTGGA AGAAATATTT
101 TTTATTAAATA AACTTCCGTC TACATTAGTG TTTCAACTAA ATACATCCTA
151 CAAAACTTTG AAGATAACGG GAAATTAAAA TAGTTGTAGG GGAAATGGAA
201 AGGAAGATAA AATCTGGGAT ATAATTTGAA GTTTTCAGT

```

Position of predicted promoter: **180** bases upstream the gfp cassette

ME22-12-07:

```

1   CTGGAGAGGA CCTATCTTAG ACCCTTATCA GCCTGGAGGC GGCACCAGCA
51  GAATCTATTA TAGGTTAACA AACTTTACTG ATTAACCTTT ATCTACCTTT
101 GATTTCCCAT GTATTTGTTT CCTCACAATT TGCTGCCCCT AGAGACTCAA
151 G

```

Position of predicted promoter: **139** bases upstream the gfp cassette

ME24-12-07:

```

1  ATATTAAGTT AAAAAGCTTC TGCACAACAA AGGATACAAT CAACAAAGTG
51  AAGAGACAAC CCACAGAACA GGAGAAAATA TTTGCAAATT ATCCATCTGG
101 CAAGGGATTA AAAACCAGAA TATATAAGGA GCTCAGACAA CTCTACAGGG
151 AAAAAAATC TAATAATCCA ATCAAAAAAC AGGCAAAAGA TTTGAATAGA
201 CATTTCTCAA AAGAAGACAT ACAAATGGCA AACAGACAGA AGAAAAATTG
251 CTGAGTATCA TTGATCTGAG GAATTACCAT GATACAAGTA GGGTATTAAG
301 GTACGATTAA ATTTGTGGAG TTTGTAGGGG CTAGGCAGGG TGATTTGGAT
351 GGGAAAGAAA CTAGAAGAAA CTTGGAAAAA TCTAGAA

```

Position of predicted promoter: **322** bases upstream the gfp cassette

MS3-12-07:

```

1  AGAGGTTACA ATAAGCTAAT GGAAGATATA ATGGTTAGTG ACTCTGTAAG
51  AAATAAGAT TAAGATCTCT TTTCTGTGC ATGAAAGCAT GATGAAAAGT
101 ATGTAGATGA GGGTGTGAGA AGTGAAGATT GGCACAGGAG AAAATTCTGG
151 GGAAGAACAA TATCTGGTGC ACCTTATGTC AATTGGGGAT AGCTGCCAG
201 ATCATTTTAT TGTGAAGAAT TCTCATTCTA CATCTAGCCT CT

```

Position of predicted promoter: **225** bases upstream the gfp cassette

MS7-12-07:

```

1  ATTCTGGCCC CCTCCGAGTG ACATAGGATG GGCTCACCAC ATGTGGTCAG
51  GCCGGCAGGG CCTCGCTGGT CCACACAATG ACCCCATAGG TCTCAAGGCC
101 TAGTGTCAGA GCAAATTCAC CGATCCTTTA CCTAACTTTC ATCATCTATT
151 TATAATTATT CCTTCAATTC TTTTCTACCA TCTCCTTCTT TTTTTTTTTT
201 TTTTTTTGAG AGAGAACTT GGCCCTGTAA GGGATTGCAA AATATGGATA
251 AAG

```

Position of predicted promoter: **121** bases upstream the gfp cassette

MS8-12-07:

```

1  CATTCTGGCC CCCTCCGAGT GACATAGGAT GGGCTCACCA CATGTGGTCA
51  GGCCGGCAGG GCCTCGCTGG TCCACACAAT GACCCCATAG GTCTCAAGGC
101 CTAGTGTCAG AGCAAATTC CCGATCCTTT ACCTAACTTT CATCATCTAT
151 TTATAATTAT TCCTTCAATT CTTTCTAAC ATCTCCTTCT TTTTTTTTTT
201 TTTTTTTTGA GAGAGCAACT TGTCCCTGTA AGGGATTGCA AAATATGGAT
251 AAAGGATC

```

Position of predicted promoter: **125** bases upstream the gfp cassette

MS9-12-07:

```

1   TAGTATTTGA TAGTGCAACA GGGTGACTAC ATCAACAATA ATTTATTATA
51  CATTATAAAA TAACTAAAGG AGCTCATGCC TGTAATCCCA GCACTTTGGG
101 AGGCTGAGGC GAGTG

```

Position of predicted promoter: **115** bases upstream the gfp cassette

MS12-12-07:

```

1   ATGGCGTGTA AGTCATCCAA TCTTGTTCTT TTTCAAAATT GTTTTGGTGA
51  TTCTAGTCCA AGGTTTCTCA ACCCTTGGGA CAATTGACCA TTTGCGCTTG
101 ATAATTATTT GCTATGGGAG GGCTGCCTGG GACACTTTAG AATGTTTAGC
151 AGCATCCCTG GCCTCTACTC ACTGGATGCT AATAGCATCT TTCCCAATTG
201 TG

```

Position of predicted promoter: **108** bases upstream the gfp cassette

MS18-12-07:

```

1   AGCCCTCTGG GAAGGGGCTG ACCAAAACCA GGCAGGGGGT TTCTGTCTAG
51  GACAAGTAGG GTCCATTTTC GGGGGAGGTG ATATATTTGC AGGGGGTGGT
101 CTGAGACCCA CAGCCAATAA ATAGGGAAGC AACATGCTTT TGGTCAGGGT
151 CCATGTGGTT AGGGTTGATT TTGATGGGGC TGGGGGGCTT AGAGTTAGAC
201 CCCCTGCCCT GGTGAGGCC CACTCCCTGG CTGGTGGGGG CGGGAGGCCT
251 CTGAAGCTGC AGCTGGAGGG GCCGGGGCAG TGTTGCATGG TGTGATTAAAC
301 CTACTAACCC AGGGCTCATC CGGGACGCTG TGGCGTGTGA CGAGGTGCCT
351 GGCCTTCTGC CAACCAACTA ACCCACTAAC CAGGCCCCGAG CATGCCCTAC
401 TAACCACCCC CCCGCGCCA

```

Position of predicted promoter: **273** bases upstream the gfp cassette

MS23-12-07:

```

1   ATGTTGATAG AGTCAATTGA TTTTGAACAG ACACATGAAA GCAAGTCAAT
51  GAGTAAAGTC TTTTAAACAA TTGGTGCTAT AACAACTGAA TAAAGGAGAG
101 AGAAAAAAAT TTCCATCTTC TTCTACACAC CATATACAGA AATTAACCTA
151 AGAGG

```

Position of predicted promoter: **145** bases upstream the gfp cassette

MS24-12-07:

```

1   ATTTTAGATG AAGAGTATCT TGAAACTTCA AGGCAGGTTG TTTTAAAAAG
51  GGTGTTGGTG GGTGTGGAG ATGGTTAATG GGTGCAAAA TATAGTTAGA
101 TAGAATGAAT AT

```

DJ1-12-07 :

```

1   AATATGAAAT AGAATTAGCC AGCAGTTCCC ACAGCTCTAA AGTGGTCTAC
51  CTGTTATTCT AAATTACAAT CACAGTAATG GTTCTGTAA AGAATACATC
101 ATCTGCTAGG TCTGGTAAAT GCATTCATCC TGATATATAT ATATATAATA
151 AGTAGTGATG GCATTGTGCC AG

```

Position of predicted promoter: **90** bases upstream the gfp cassette

DJ4-12-07:

```

1   TGGTGGGAGG TAATTGAATC ATAGGAGCAG TTACCTCCAT GCTGTTCTTG
51  TGATAGTGAG TGAATTCTCA TGAGTTTGAT GGTTTTGTAA GGGGCTCCCC
101 GCCGCTTCAC TCTGCACTTC TCCTTGCTGC CACCATGTGA AGAAGGATGT
151 GTTTGCTTCC CCTCCACCA TGATTGTAAG TTCCTGAGG CCTCCCAGCC
201 CTGCAGAACT GTGAGTCAGT AAAACCTCTT TCCTTTTATA AATTACCCAG
251 TCTCAGGTAT GTCCTTATAG CAGCATGAGA ACGGACTAAT ACATCTCCTT
301 AGGCTCCTTT TAGCTGTGAT AATTTCTCAG ACTTCTCTTG TTTTGGTGA
351 CCTTGATAGT TTTGAGGAGT ACTCTTCAGA CATTTTGTAG AATGTCCCTC
401 AATTGAGATT TGTCTGTTAT GTTCTCATTT ATTATACTGG GATTATGGGT
451 TTGGGGGAAG AAGAAGACAG TGGTAAAGTG GTATCTT

```

Position of predicted promoter: **98** bases upstream the gfp cassette

DJ5-12-07:

```

1   TGATGAATCC CCTAATGATT TTGGTAAAA TCATTAAGTT AAGGTGGATA
51  CACATCTTGT CATATGATCA AATGGTTTCG CGAAAAATCA ATAATCAGAC
101 AACAAAGATGT GCGAACTCGA TATTTTACAC GACTCTCTTT ACCAATTCTG
151 CCCCGAATTA CACTTAAAC GACTCAACAG CTTAACGTTG GCTTGCCACG
201 CATTACTTGA CTGTAAACT CTCACTCTTA CCGAACTTGG CCGTAACCTG
251 CCAACCAAAG CGAGAACAAA ACATAACATC AAACGAATCG ACCGATTGTT
301 AGGTAATCGT CACCTCCACA AAGAGCGACT CGCTGTATAC CGTTGGCATG
351 CTAGCTTTAT CTGTTGCGGC AATACGATGC CCATTGTACT TGTTGACTGG
401 TCTGATATTC GTGAGCAAAA ACGACTTATG GTATTGCGAG CTTCAGTCGC
451 ACTACACGGT CGTTCGTGTA CTCTTTATGA GAAAGCGTTC CCGCTTTCAG
501 AGCAATGTTC AAAGAAAGCT CATGACCAAT TTCTAGCCGA CCTTGCGAGC
551 ATTCTACCGA GTAACACCAC ACCGCTCATT GTCAGTGATG CTGGCTTTAA
601 AGTGCCATGG TATAAATCCG TTGAGAAGCT GGGTTGGTAC TGGTTAAGTC
651 GAGTAAGAGG AAAAGTACAA TATGCAGACC TAGGAGCGGA AAAGTGGAAA
701 CCTATCAGCA ACTTACATGA TATGTCATCT AGTCACTCAA AGACTTTAGC
751 TATAAGA

```

Position of predicted promoter: **332** bases upstream the gfp cassette

DJ7-12-07:

```

1   AAAATAACAC GAATAACGTG GCAGTGTCAT GAGTAAACAT AAAAAGATAA
51  ACTGGTGTGT CTGCATGTCA CGGCTGGCAT AGTGGGCATG AGATCCATGC
101 CCAGTGCTCA TAAATAGTCT GTTCACATCC TGCTTTCTGT TCCAATCTTT
151 TCGTCTCTCC TTTTGGTTGT TACAACTG

```

Position of predicted promoter: **174** bases upstream the gfp cassette

DJ8-12-07:

```

1   AATGTCTACA TGGTACATTG TACATAATTT AATTTTCCCA GTTGAACATC
51  GTGACACTAT ATGAACCATG AATGCACGAA ATAGGAATTC CAAATGTCTT
101 CCATATGGTA AGAGGACTTT TGGATACTTC TGCAGGAAGC TTAAGAGTAG
151 CTCGAAATTA AGTTATGCCT GTTCTCAGTG ACTTTTATTA TGTATGCCTA
201 AGGTATTAGC ACTTTTCTCT GAGTGGGACC TTGAGGTGAT AGATAGATGA
251 GTCTGCTGTC TGTATTTAGG CACTTGGCTC ATTTCTTGAG AGCTGTGGAA
301 GGCATGTGGA TGCTCATTAA GTTTCCAACC ATTCCCCGAG AGCAACAGCT
351 GGTGAATCAA AGCTTCCCTG GTATGTCATG ACTGGTTTCC TCAAAGTCTT
401 GGAAATAATG GATGGCATTC ACATGCTCTC TGTAGCACCA ACAGATAAGA
451 AGGAGATGAA TAGCAACAGG AAGCATTTC AAAGCATAAA GAAACACACA
501 GGTCTAGGAA GATTATGAGT AATGCGCTTG CTACCACACA TTCTATATCC
551 ATATATGGTG GTTTCATCTT TTAAATACTG CTACTCGTTC AAAGTATGAA
601 ATCATAGTTA ACTGTCCTGT GGTGGCTACC TAATAGAAGA AATGTAAATT
651 AAAGGAAAAA AGAACATTCT AAGTTAAGAA AACTATGTCA TAAAATTTAA
701 GCCTTTAGCT CTACATTTAT AAATATCCAA TGACCACTAA TAGTTGTGAC
751 GTATCTTTAA ATTGTCT

```

Position of predicted promoter: **120** bases upstream the gfp cassette

DJ12-12-07:

```

1   ATAAAGTGTG TATGTATTTA TATAAATTAT GTAAAAGCTA GTAGGTATTT
51  TACAAAATTT AGATTTTCCT TTCAATGAAT AGGAAGTCCT TGGTAGGTCT
101 T

```

Position of predicted promoter: **92** bases upstream the gfp cassette

DJ18-12-07:

```

1   AAATCAGGGT AATCAGCATA CCCACCATCT CAAACACTTA TTATTTCTTT
51  GTGGTGAGAA CATTCAAAAT CCTCTCTTCT AGCTATTTTG AAATATACAA
101 TACATTATTG TTAACTATAG TCACCCTACT GTGTAATAGA ACTCCAGAAG
151 GTATTCCTCC TCTCTAACTG TAACATAGTA TCCATTG

```

Position of predicted promoter: **99** bases upstream the gfp cassette

DJ22-12-07:

```

1  ACCAAGGTCA CGGACTCTTG TCTTGGGTCT GCCTATGGGG GCATTGTAAC
51 ATACCTCTGC ACTG

```

DJ6-12-07 :

```

1  AGAGATGTTT AACACACTCC TGGTCACTCG ATGAGCAGCT GGCAGAACGG
51  GAGCCCGTAT CTGTGGTTTG GAGATCATCT CTGTATCTAA TTTGTATTAT
101 GACTCTTCTT ATTGTAAC TA TTTCTTTAT TATACTGAAA CACTTTGTGC
151 CTTCACTCTC TTGCCTTGGC ACCTAGGTAA TCTTAGGTCC ACACTTTGGG
201 AAGCCAAGGC AGGT

```

Position of predicted promoter: **163** bases upstream the gfp cassette

DJ19-12-07 :

```

1  ATATTTGGAA CTTCTGTAGA CATTTCAAAC TCACAGATCC TTTCTAACAT
51  TGGGATACCC AGCTAAGGAA TGAGGATCAG GCCTCTGGGG ACCCCATGAG
101 GCTAAAAGAT ATCAAGGACA GACTTGAAAT TCTAGATTTTT ACAATACAAC
151 AAGAGATTGA TCATGTATCC CTTCTAGGCC TGGCCCATAA AAACCTCCTA
201 TGCATG

```

Position of predicted promoter: **110** bases upstream the gfp cassette

DJ23-12-07:

```

1  CACCTGCAGC TTGATAGAAT ATGGCAAGGT GTTATATAGA AATATCTTCT
51  TACGTCTTTC TTGCTCTTCT TCAATAACAT AATAACAATA AAAAGACAAA
101 CTGTAATGTG TTATGTACTT GTTGTACCCA GCATTGGTCT AAGTGCCTTA
151 TAGGTATTGT CCCTAAAATC CTCTAAAAAT TGATCAAATA AAAACCATTA
201 TCATTTTTAAT GTAATGAGAA GAAATTCTGG TTCAAAGTGA CTCATAAAT
251 CCCTTGAGTT AACCAGGCTA ATAGGAGGCT CAGCTCAAAC TCAAAC TCAA
301 ACTCAAAC TC AGGTATATTA AATACAATGG TGATATAGAA GGAA

```

Position of predicted promoter: **189** bases upstream the gfp cassette

JB6-12-07:

```

1  TTTTAAAATA ACTCATTGCA GAGAACGAGG TGCTCTTTGA CTGTCAGCGG
51  ACAGAGTTCA CAGATTCCTT GAGGGCTGGA TGGATGAGAT ACAGAAAGAC
101 AGAGCTAGGA TCCAATGCCT CAAAGAATAT TACAATAAAA AATAGATTTT
151 AGGAAATGCT AGAGTAGACA GTAAGCTCCT TTGAGAATAA AAAAAAATT
201 CTAATTCACT TCTGCATTCC TGTTGCCTAG CATAGTGTGT CAATTTATTC
251 TGGTGCATTG TTAAGTACAT AACATAGCTC ATAAGTTTAG AGGGATAGAC
301 CTGTACCCCT CTGTGTGTCC AATGACTACA AGCCAAATTG AAGGTGTAGA
351 AAATAGTCAC TCTTATAAAA GCAGTGCAAC AAGAAAAACA TAAACCATAC
401 TGAACATAAG CCTATAAAA

```

Position of predicted promoter: **235** bases upstream the gfp cassette

JB7-12-07:

```

1  ACATCCCGTT AAAAAGCTTC TGCTCAGAAC AGGAAACAAC GTGAGGAGAC
51  AACCACAGA ATGGGAGAAA ATATTTGCTA ACTGTCCATC TGACAAGGGG
101 TTAACAATA AAATATATAT CCGGAGCTCA AGCCACTCTA TGGAAAAATC
151 TACTATTCTG ATTTAAAAAT GGGGGAAGA TCTCTATGGA GAAAATCAGA
201 AACCTGATT GTGAGACAAA GGGGACATTG TAACGAGGAA GAGGTACCGC
251 GTCCGTGGAA TAACATTAAA AGAAGTCCAC TTTCTCCATG TGATTTATAA

```

Position of predicted promoter: **185** bases upstream the gfp cassette

JB8-12-07:

```

1  TCTATAGTCA CCCTGACGGA TGCAAGGGAC GGCTGCTCGA GTTCTATAGT
51  GTCCCGTAAA TCTTATGAGT ATGAAACATA AGGTTATGTA TTAATTGCAC
101 CCTCGTTCAT TCTCTTATTT GTTTATATGG GGCACCTATA CTACAATCTG
151 CTCTGATGCC GCATAGTTAT TCCAGCCGCG ACACCCGCCA ACACCCGCTG
201 AAGCGCCCTG ACGGGCTTGT CTGCTCCCGG CATCCACTTA CAGACGCGCT
251 CTGACCGTCT CCGGGAGCTG CATGTGTCAC A

```

Position of predicted promoter: **235** bases upstream the gfp cassette

JB9-12-07:

```

1  TGACAATAAA TGGTGCTGGA ACAATTGGAC ATACACATGC AAAAAAAAAA
51  GACCCAAACA CTTATCAATA CTTTACACCA ATAAAAAAT TAACCTCAAA
101 TTGATCCTCC CACTTCAGCC TCCTGAATAG CTAGGACAAC AGGTGCATGA
151 CATGAGACCC GGCCAATTTT TTAATTTTTT GTGGAGACGG TGAGGGGGGT
201 GGGGTGTCTC ACCATGTTGC CCAGGCTGGT CTTGAACTCC CCACCTCAAG
251 T

```

Position of predicted promoter: **199** bases upstream the gfp cassette

JB10-12-07:

```

1  CCTCATAATT TCCCCAAATC ATAACCATGT GTGAATTTTT TTTGAGTTGA
51  GCAGGTTGCT ACCGCCAATG AGTCTTCCCA TAGGGGTGAT GAGACTGAAT
101 GCCGCTCAAT TTCCCTGCCG GACCGAACTG TTAACAGATC CTCTCGATTT
151 CAGATTGTGG CACGCTGGCT TTGCCTGTCC CAATCTTGTT TCGGAACATG
201 AACCCCCGCA GGCCAACATG TTAGGTGAAG CATTAAGGCG ATTAGCGTTC
251 TGATATCAAA ACGCTCTGAG CTGCTCGTTC GGCTATGGCG GCTCCTATGC
301 CGTAGGCAGG ACTTTTCACT CTCGGAAGGT TTCTTCGCTG ACCATTCCCT
351 ACGGACATGT CTGAACCAGG TGTTTGTTG TTCAAATTTG GTCACGGACT

```

Position of predicted promoter: **230** bases upstream the gfp cassette

JB11-12-07:

```

1  TACATGTCTG AGATGGCACT AGGTCTACCA TACACAGCTA AGTAAATAAA
51 TGCTACAATA CAGTAATATA ATAG

```

JB12-12-07:

```

1  TATACCTTAA CAAGTATTCT CATCAAAAAC CACTGGAGGC CTGGCCCAGT
51 GGCTCATACC TATAATCCCA GCACTTTGGG AGGCCAGTGT GGGAAAG

```

Position of predicted promoter: **79** bases upstream the gfp cassette

JB13-12-07:

```

1  TATTTATTTA AGCCTTCTAT GTACTCTTTC CACAGTGTCT CTTGATTTCT
51 AGTGCCTTGG ATTTGAAAG

```

JB16-12-07:

```

1  TGTGTTTTTA AAAAGCCTTT GATGCGATTT TGATCCACCT GTCCCGGCCT
51 CCCAAAGGGT TGAATTATA GGCATGAACC ACCACGTCCA GCCTGATTTT
101 TTATTATTAC AAACAGTACT GCAGTAAACT ACTTTTTTTA CATGTATCTT
151 TGTGTATTTG TGTATTTCTG CAGGATAAAT TTATAAAAGT AGAAACAGTA

```

Position of predicted promoter: **99** bases upstream the gfp cassette

JB17-12-07:

```

1  GGTGATCCAC CCACCTCAGC CTCCCACAGT GCTGGGATTA CAGGCGTGAG
51 ACACCATGCC TGGCCCACTT TCCTACTTCT TTGATGGCAA CTCCATCTTT
101 CCTGTTGCTC AAGTCAGTTA TCCCTAATGC CACCTTCTCT CTAACAACAAC
151 AAATGCTGTT GGCTTTACTT GCAGAATATA TCCAGATTCT AACCACGTCT
201 TAACATTACT ACCACCTGGC CCAAGCTTTC ACCTACATTA CTGCAATATT
251 TCTTCAGTAG ACTTCTCAGA TTCTACACTA TCCTCACCAG CCACCCAAGA
301 G

```

Position of predicted promoter: **181** bases upstream the gfp cassette

JB18-12-07:

```

1  GCCTGCCTCG GCCTTCCAAA GTGCTGGGAT TATAGGCGTG AACCACTGTG
51 CCTGGCCTGA ATTGAATTTT TGAAATGAGC AATATATTTT CTTTTACTTT
101 TCATATAAAA GTTAGATGAA TCACATTATC TTTCATGGCC AATTTAATTG
151 CAAGATA

```

Position of predicted promoter: **118** bases upstream the gfp cassette

JB19-12-07:

```

1   AAACTTATAA AAGTATTGCA TCAATGACAA ATGCATTGCC GTTTTATTGA
51  AGTTGATAAC AGAATTGTAG TTATAAACAG TCTTTTAGTA TACTGTCTTT
101 GTCCTTG

```

Position of predicted promoter: 79 bases upstream the gfp cassette

JB23-12-07:

```

1   ACAATGGTAT GAAGCTAGAA ATCGATAACA GCAAGAGTTT ATAAAATTTA
51  CAAATATGGA AAATTAAACA ACATGTTTCAT AAACAACCAA TGGATCTTAA
101 CATTTTCAGTT AACCACGCA ACAAGTAATG CAGCTGTGTG ATTTTGAGAC
151 AAGTTACTTA GCCCCTCTGT GCCTCAGTTT CCTTATCTGT TCATAAAACA
201 AGTATATTAG ACTAGAGATT ATTTTCATTC CTAAATTTTG TGATGTTATA
251 GGGCTGTTCC TGAGTGAAAT TTCACAGAGG AAAGGATGTA CCTCTCAATG
301 GCTGCCATAG CTCATATTGT GGCTTTTATG TTTGAAAAGA GACTTTCATT
351 TTGGAAGGTG GGCCTGGTGG TTGTGATGTC CTGCAGTTTG AAGGAACAAG
401 GCTAACAATC AGAGGGCCTG AATTAGGTAC TGGCTCCACT ATTTAGTGGC
451 CGGAAAACCT TACAGGCAGA ATCATTCCCT GGAAATAAAG CTCCCCTGGA
501 TACTATCTGT TCTGTATGAG CATCTGTTCC TCTCCCATGA AATGGACATA
551 ACTGTCAGTA CCTTATCTGT TGACTTCCCA AGGTGGAGTG AGAATCACAT
601 TAAATCATTT ACTCAGTCAG CTTGCACTAT GCAGAAGATT TTTGGCATTT
651 TGGTGGTATA AACCATAGGT GTGCATATAG CTCAAAAGAT TCTATGCTGT
701 GATTCTAATG AACTCTCAAC TCTTTAGATC ATCTAATTTA AACAAATGTCC
751 TGCAAGCAAG CGGGAGTTAA CTATTCCTCT TTTTGT

```

Position of predicted promoter: 137 bases upstream the gfp cassette

JB24-12-07:

```

1   TTTTGGAGTT CTACCCTACA CATGTGTGTC TTAATATTCA GCAAAGACTC
51  AAAGGAAGTG CTGTGTAAAT TTCTTACTTT CATTTTTTTG TAAAGTTCCT
101 TCTTCCCAG AACTCTGCCC TGCAGCTTCC AGTCCTCCAA GACTCTCTGA
151 ACTTGGATCT TGACACCTTT CTGGAATTCC TTTTGGTTTA AAATCTTTCC
201 ATGTTGTCTG TGAACAGATC TTGGAAATAT CGTTTCTCAA ACCTATGAAT
251 TTCACAGTCA AGTTCTACTT TACATCTCTT GTTCTTAGGC TCCCATGCTG
301 GAAGATAGTC ACACACTTTT TTTCTTCAGT GATGATCTTC ATTTTAGCAG
351 CTTGTTGGCA TTGATGGACC AGAAGAACTCT AAGAGAGTTA TAACCATAAG
401 TGTATGATAA TGATCTCCAT GCAACAGAGC AAATTTTTAT GTGCTGACAT
451 CACATCAGGA CGCTCTTCTG CCCACTGTCT TCCAGTGAAG TCCCCTGTG
501 TCTTTCACTC AGAGTGAAAG CCTAGGTTAC CAGGGTCTAG GGTCCCCCTA
551 CTCTCCAGG TTCTTCCCCC TTCAAGACTT TTGTGCAGGT

```

Position of predicted promoter: 231 bases upstream the gfp cassette

Inserted sequences in GFP positives clones and position of predicted bacterial promoters near the TSS

*Nucleotides in light grey represent the -10 and -35 predicted bacterial consensus motifs

Position of predicted promoter was calculated from the final -10 nucleotide to the beginning of the GFP cassette (TSS)